

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
12.09.2001 Bulletin 2001/37

(51) Int Cl.7: C12N 15/52, C12N 9/00

(21) Application number: 01104583.8

(22) Date of filing: 06.03.2001

(84) Designated Contracting States:  
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU**  
**MC NL PT SE TR**  
Designated Extension States:  
**AL LT LV MK RO SI**

- Frey, Bruno Dr.  
82377 Penzberg (DE)
- Antranikian, Garabed Prof. Dr.  
21218 Seevetal-Hittfeld (DE)
- Boehlke, Kristina  
22761 Hamburg (DE)
- Pisanl, Francesca Marla Dr.  
00140 Rom (IT)
- Rossi, Mosè Prof. Dr.  
80072 Arcofelice/Napoli (IT)

(30) Priority: 11.03.2000 EP 00105155

(71) Applicant: Roche Diagnostics GmbH  
68298 Mannheim (DE)

(72) Inventors:  
• Sobek, Harald Dr.  
82377 Penzberg (DE)

(54) **Mutant B-type DNA polymerases exhibiting improved performance in PCR**

(57) Claimed are thermostable mutants of B-type DNA polymerases having a Y-GG/A amino acid motif between the N-terminal 3'-5'-exonuclease domain and

the C-terminal polymerase domain whereas the tyrosine of the Y-GG/A amino acid motif is mutated and whereas these mutant DNA polymerases are suitable for PCR.

	E	R	R	R		G(Y)	K	E	E	L	W	E		consensus							
<i>T. aggregans</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>T. litorea</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>T. lummodans</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>T. spec. 9N7</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>T. gorgonensis</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>P. spec. KOD</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>P. abyssal</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>P. furiosus</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>P. horikoshii</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>M. jannaschii</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>M. caldococcus</i>	V	R	R	L	T	T	V	L	G	V	K	R	G	L	W	E	N				
<i>S. solfataricus</i>	T	S	A	L	I	K	G	K	V	K	A	V	I	D	P	A	G	I	F	F	N
<i>S. acidocaldarius</i>	T	A	A	V	I	K	G	K	V	K	A	V	I	D	P	A	G	I	F	F	N
<i>P. islandicum</i>	T	A	A	V	I	K	G	K	V	K	A	V	I	D	P	A	G	I	F	F	N
<i>P. coccolitum</i>	S	E	A	L	I	K	G	K	V	K	A	V	I	D	P	A	G	I	F	F	N
<i>A. parvum</i>	V	G	A	L	I	K	G	K	V	K	A	V	I	D	P	A	G	I	F	F	N
<i>S. chaetoceros</i>	T	A	A	V	I	K	G	K	V	K	A	V	I	D	P	A	G	I	F	F	N

SEQ. ID. NO: 16  
SEQ. ID. NO: 17  
SEQ. ID. NO: 18  
SEQ. ID. NO: 19  
SEQ. ID. NO: 20  
SEQ. ID. NO: 21  
SEQ. ID. NO: 22  
SEQ. ID. NO: 23  
SEQ. ID. NO: 24  
SEQ. ID. NO: 25  
SEQ. ID. NO: 26  
SEQ. ID. NO: 27  
SEQ. ID. NO: 28  
SEQ. ID. NO: 29  
SEQ. ID. NO: 30  
SEQ. ID. NO: 31  
SEQ. ID. NO: 32

Figure 8

## Description

## FIELD OF THE INVENTION

5 [0001] Subject of the invention is a thermostable mutant B-type DNA-polymerase having a Y-GG/A amino acid motif between the N-terminal 3'-5'-exonuclease domain and the C-terminal polymerase domain in the wild type form whereas amino acids of this motif are substituted in the mutant form of the DNA polymerase and whereas these mutant DNA polymerases are suitable for PCR reactions. Thermostable mutants according to the present invention exhibit better performance in PCR reactions compared to the wild type DNA polymerase. A further embodiment of the present invention is the use of these thermostable mutants of the B-type DNA polymerase for polymerase chain reactions (PCR) and other nucleic acid synthesizing reactions. Another subject of the present invention is a method of producing the inventive mutants, vectors and cell lines comprising genes encoding the inventive mutants.

## BACKGROUND OF THE INVENTION

15 [0002] DNA-dependent DNA polymerases containing proof-reading activity have to coordinate two catalytic activities: the DNA polymerase activity and the exonuclease activity. For polymerase I type DNA polymerases (*E. coli* Pol I) as well as for B-type DNA polymerases, these catalytic activities are located on structurally distinct protein domains (Truniger, V., Lázaro, J., Salas, M. and Blanco, L. (1996) *EMBO J.*, 15(13), 3430-3441; Pisani, F. M., De Felice, M. and Rossi, M. (1998) *Biochemistry*, 37(42), 15005-15012). In B-type (eukaryotic-type) DNA polymerases, the coordination of the two catalytic activities was proposed to take place intramolecularly in the conserved motif Y-GG/A located between the N-terminal 3'-5' exonuclease and the C-terminal polymerization domain (Truniger, V., Lázaro, J., Salas, M. and Blanco, L. (1996) *EMBO J.*, 15(13), 3430-3441; Pisani, F. M., De Felice, M. and Rossi, M. (1998) *Biochemistry*, 37(42), 15005-15012). For the Klenow fragment of *E. coli* DNA polymerase it was described, that the editing can be an intermolecular or intramolecular process involving dissociation and reassociation of the DNA depending on the local context (Joyce, C. M. (1989) *JBC*, 264(18), 10858-10866).

Truniger et al. (Truniger, V., Lázaro, J., Salas, M. and Blanco, L. (1996) *EMBO J.*, 15(13), 3430-3441) demonstrated for the mesophile replicative DNA polymerase of bacteriophage  $\phi 29$  that mutations in the Y-GG/A motif can lead to phenotypes favoring either polymerisation or exonucleolysis compared to the wild type enzyme. They could show that this effect is related to altered (ss)DNA binding parameters and that the motif is important for the communication between the polymerase and exonuclease active site in a combination of structural and functional roles.

30 [0003] For the DNA polymerase of the thermophilic crenarchaeon *Sulfolobus solfataricus* (Sso) a region of 70 amino acids (region 1) involved in enzyme-DNA interaction was determined (Pisani, F. M., Manco, G., Carratore, V. and Rossi, M. (1996) *Biochemistry*, 35, 9158-9166). It is located in the connecting part between the exonuclease domain and the polymerase domain and contains the Y-GG/A motif. By mutational analysis of the amino acids in the Y-GG/A motif, it could be shown that the amino acids in this part of the enzyme determine the processivity of the proofreading function (Pisani, F. M., De Felice, M. and Rossi, M. (1998) *Biochemistry*, 37(42), 15005-15012). Based on the crystal structure of bacteriophage RB69 DNA polymerase, Truniger et al. proposed a direct interaction of the tyrosine with the phosphodiester bond between the two nucleotides preceding the one acting as template (Truniger, V., Blanco, L. and Salas, M. (1999) *J. Mol. Biol.*, 286, 57-69).

## DESCRIPTION OF THE INVENTION

45 [0004] The subject of the present invention was to provide thermostable DNA polymerases exhibiting an improved performance in PCR. Especially, thermostable mutants of a B-type DNA polymerase are provided which exhibit improved PCR performance. The inventive mutants of the B-type DNA polymerase have mutations in the Y-GG/A amino acid motif. Preferred mutations refer to the position of the tyrosine in the Y-GG/A amino acid motif. Other mutations affecting the motif could also influence the performance of B-type DNA polymerases in PCR.

50 [0005] According to the present invention an improved performance of a DNA polymerase in PCR is defined as a performance that results in higher yields of PCR product, or the amplification of longer DNA targets. Additionally, improved PCR performance can be defined as improved fidelity during the amplification process.

[0006] Preferred mutant B-type DNA polymerases have mutations at the position of the tyrosine in the Y-GG/A amino acid motif. Preferred mutants of B-type DNA polymerases according to the present invention have phenylalanine, tryptophan or histidine at the position of the tyrosine. Other preferred mutants of B-type DNA polymerases according to the present invention have asparagine or serine at the position of the tyrosine. These mutant polymerases described here, in which the tyrosine of the Y-GG/A motif was substituted, exhibit an improved performance in PCR.

55 [0007] In a preferred embodiment the inventive mutant B-type DNA polymerase is a mutant of a B-type DNA polymerase obtainable from Euryarchaea, more preferably from *Thermococcus aggregans* (Tag). Especially preferred is a

mutant of a B-type DNA polymerase from *Tag* of about 94 kDa size with a temperature optimum of  $\geq 80^\circ \text{C}$  and the ability to perform polymerase chain reactions.

[0008] The present invention is described in detail for the B-type DNA polymerase from *Thermococcus aggregans*, but the invention could also be applied to other B-type DNA polymerases. Preferably to those B-type DNA polymerases showing a high degree of homology ( $\geq 80\%$ ) to the DNA polymerase from *Thermococcus aggregans*.

The B-type DNA polymerase from *Thermococcus aggregans* exhibits a high degree of amino acid sequence homology to B-type DNA polymerases of other *Thermococcus* species. The homology of the DNA polymerases was calculated using the program Blast 2 (Tatusova, T.A. and Madden, T.L. (1999) *FEMS Microbiol. Lett.* 174, 247 - 250). The homology of the B-type DNA polymerase from *Thermococcus aggregans* to the homologue enzymes from *Thermococcus* species is: 93% (*T. litoralis*), 87% (*T. gorgonarius*), 86% (*T. furiosus*) and 87% (*T. spec.* 9N7). The homology of the *Tag* DNA polymerase to polymerases from *Pyrococcus* species is: 86% (*P. abyssii*), 86% (*P. horikoshii*), 86% (*P. spec.* KOD) and 85% (*P. furiosus*). A lower homology is calculated to other B-type DNA polymerase from different euryarchaeota: 59% (*Methanococcus jannaschii*), 56% (*Methanococcus voltae*), 51% (*Methanobacterium thermoautotrophicum*) and 56% (*Archaeoglobus fulgidus*). To B-type DNA polymerases from crenarchaeota and bacteriophages the homology is found as follows: 46% (*Sulfolobus solfataricus*), 42% (*Sulfolobus acidocaldarius*), 41% (*Sulfurisphaera ohwakuensis*), 51% (*Aeropyrum pernix*), 40% (*Pyrodicticum occultum*), 43% (*Cenarchaeum symbiosum*), 38% (bacteriophage T4) and 39% (bacteriophage RB69).

[0009] As described above several mutations in the Y-GG/A motif were performed for the *Sulfolobus solfataricus* (*Sso*) and the  $\phi 29$  DNA polymerases. The observed effects on polymerase activity (pol) and exonuclease activity (exo) of these mutations do not completely correspond to the effects obtained for the *Tag* DNA polymerase. Thus the effect of the mutations on the performance of the mutants in PCR was not predictable.

The mutant Y387F of the *Tag* DNA polymerase exhibits a higher pol/exo ratio compared to the wild type *Tag* DNA polymerase. Similar results were described for *Sso* and  $\phi 29$  DNA polymerase. The mutant G389A displays the opposite effect than the corresponding mutant in  $\phi 29$  DNA polymerase: while G $\rightarrow$ A in *Tag* DNA polymerase almost knocks out polymerase activity, in  $\phi 29$  DNA polymerase G $\rightarrow$ A mutant this activity is clearly enhanced. For mutants of the *Sso* DNA polymerase a change of exonuclease processivity was described. Again, this was not observed for mutants of the B-type *Tag* DNA polymerases. Thus, a prediction of the effect of analogous mutants in the Y-GG/A motif could not be made.

[0010] In summary, although it has been described in the prior art that the Y-GG/A motif plays a role in the coordination of the DNA polymerase activity and the exonuclease activity, the observed changes of the pol/exo ratio of the prior art DNA polymerases do not strictly correlate to the changes observed for the inventive mutants of the *Tag* DNA polymerase. Furthermore, it has not been described that the Y-GG/A motif is important for the performance of B-type DNA polymerase in PCR. Additionally, there is no correlation between the changes of the pol/exo ratio and the improvement of the performance of DNA polymerases in PCR. For instance, the mutant Y387H does not exhibit a change of pol/exo ratio compared to the wild-type, but it exhibits improved performance in PCR. Furthermore, a significant enhancement of fidelity was observed for the mutants Y387N and Y387S of *Tag* DNA polymerase.

[0011] Results obtained for the mutants of *Tag* DNA polymerase are described in more detail below.

#### Enzymatic activities of wild type and mutant *Tag* DNA polymerases

[0012] The enzymatic activities of the wild type enzyme and the mutants of *Tag* DNA polymerase were determined and analyzed (Figure 1). The DNA polymerase activity was determined as described in Example 2. According to the effect of the mutations on the polymerase activity three groups of mutants were defined: i) mutants with enhanced DNA polymerase activity (mutant Y387F), ii) mutants having a similar or slightly reduced DNA polymerase activity compared to the wild type (mutants Y387W and Y387H) and iii) mutants with reduced DNA polymerase activity (mutants Y387N, Y387S, G389A).

The exonuclease activity was determined as described in Example 3. According to the effect of the mutations on the exonuclease activity two groups of mutants were defined: i) mutants with similar exonucleolytic activity as the wild type enzyme (mutants Y387F, Y387W, Y387H), ii) mutants with enhanced exonuclease activity (mutants Y387N, Y387S, G389A) in comparison to the wild type enzyme.

[0013] From the data obtained for polymerase activity and exonuclease activity the ratios of both activities (pol/exo) were calculated for the wild type enzyme and the mutants of *Tag* DNA polymerase (Figure 1). Three mutants showed a higher or similar pol/exo ratio as the wild type enzyme (mutants Y387F, Y387W, Y387H). Three mutants showed a clearly reduced pol/exo ratio in comparison to the wild type enzyme (mutants Y387N, Y387S, G389A).

#### PCR performance

[0014] Wild type and mutant enzymes were submitted to polymerase chain reactions on lambda DNA in a buffer optimized for this purpose. All mutants except for mutant G389A were able to perform PCR, but yielded different

amounts of product with a constant amount of enzyme (1 pmol). With increasing length of the DNA target, differences in performance of the enzyme were shown (Fig. 2). With 1 pmol of the mutants Y387S, Y387N and G389A no PCR product could be obtained for the amplification of a 3.3 kb fragment. 1 pmol of the wild type DNA polymerase could not amplify fragments of 5.0 kb length. The mutants Y387W, Y387F and Y387H were able to amplify a fragment of 7.5 kb length. As control *Taq* DNA polymerase, *Pwo* DNA polymerase and Expand™ High Fidelity PCR System (Roche Molecular Biochemicals) were used.

[0015] The differences in PCR performance were also shown by the amplification of a 2 kb fragment applying different elongation times in the PCR runs. Under these conditions, all enzymes tested except the mutant G389A were able to amplify a 2 kb fragment at an elongation time of 90 sec/cycle. The mutants Y387F, Y387W and Y387H were able to amplify the fragment at a reduced elongation time of 40 sec/cycle. The mutant Y387H was able to amplify the target in a elongation time of 30 sec/cycle (Fig. 3).

#### Exonuclease processivity

[0016] The exonuclease processivity of the enzymes was studied in an experiment based on the heparin trap method (Reddy, M. K., Weitzel, S. E. and von Hippel, P. H. (1992) *J. Biol. Chem.*, 267(20), 14157-14166; Pisani, F. M., De Felice, M. and Rossi, M. (1998) *Biochemistry*, 37(42), 15005-15012). A constant amount (1 pmol) of *Tag* DNA polymerase or its mutants was incubated for 4 minutes at 68 °C with a 5'-DIG-labelled 24mer oligonucleotide in the absence of nucleotides. In the absence of heparin, the oligonucleotide was continually degraded by the *Tag* enzymes (positive control, Fig. 4, lanes "-"). The function of the heparin trap method was demonstrated by addition of heparin and MnCl<sub>2</sub> before the binding of the enzyme (negative control, Figure 4, lanes B). Single turnover conditions (addition of heparin and MnCl<sub>2</sub> to start the reaction after the binding of enzyme) resulted in exonucleolytic degradation of the oligonucleotide by the *Tag* DNA polymerases (Fig. 4, lanes A). The enzymes showed differences in the exonucleolytic activity as shown by the different amounts of remaining oligonucleotide that was not degraded. However, for all enzymes tested the oligonucleotide was degraded to a similar extent (8 nt). This indicates a similar exonuclease processivity for the enzymes.

The *Thermococcus gorgonarius* DNA polymerase, which exhibits a strong exonuclease activity, was used as a positive control. It degraded the 24mer oligonucleotide in the absence of heparin to oligonucleotides of less than 15 bases length (Figure 4, lane "-"). Under single turnover conditions a strong degradation (11 nt) of the oligonucleotide is observed (Figure 4, lane "A").

#### Fidelity

[0017] The error rates in amplification were determined for the mutant enzymes and the wild type DNA polymerase. The PCR-based fidelity assay described by Frey and Suppmann (Frey, M. and Suppmann, B. (1995) *Biochemica*, 2, 34-35) was used. This method is based on the amplification, circulation and transformation of the pUC19 derivative pUCQ17, which contains a functional *lacI*<sup>q</sup> allele (Barnes, W. M. (1994) *Proc. Natl. Acad. Sci. USA*, 91, 2216-2220). PCR-derived mutations in *lacI* result in a de-repression of the expression of *lacZα* and subsequent formation of a functional β-galactosidase enzyme, which can be detected on X-Gal indicator plates.

[0018] In five independent runs, a mean error rate of  $5.0 \times 10^{-6}$  was found for the wild type *Tag* DNA polymerase. This value is in between the mean error rates of  $1.8 \times 10^{-6}$  for Expand™ High Fidelity PCR System (Roche Molecular Biochemicals) and  $1.3 \times 10^{-5}$  for *Taq* DNA polymerase (Roche Molecular Biochemicals) determined in the corresponding experiments. For a better comparison of the data, we plotted the quotient of the error rate determined for *Taq* DNA polymerase divided by the error rates determined for the *Tag* DNA polymerase and its mutants. In the independent experiments the error rate for *Taq* DNA polymerase varied from 1.2 to  $3.05 \times 10^{-5}$ .

Figure 5 shows the quotients of the error rates of the wild type enzyme and mutants of *Tag* DNA polymerase. The error rates of the mutants showing improved PCR performance (Y387W, Y387F, Y387H) did not significantly differ from the values obtained for the wild type enzyme. The mutants with enhanced exonuclease activity (Y387N, Y387S) showed improved fidelity rates (Figure 5). For the mutants Y387N and Y387S mean error rates of  $6.3 \times 10^{-7}$  and  $6.2 \times 10^{-7}$  were determined.

[0019] In contrast to the φ29 DNA polymerase and the *Sso* DNA polymerase, the *Tag* DNA polymerase, is suited for PCR. The mutant enzymes (Y387F, Y387W, Y387H) with an aromatic amino acid in the position of the tyrosine showed a similar or only slightly enhanced DNA polymerase activity (mutants Y387F, Y387W, Y387H) but an improvement in PCR performance.

[0020] In the fidelity assay it was found that the mutants Y387F, Y387W and Y387H showed no significant change in their error rate. By contrast, the mutants Y387N or Y387S showed higher exonuclease activity and displayed an improved fidelity.

[0021] Subject of the present invention is also a method of producing the inventive B-type mutants comprising the

following steps: cloning and mutagenesis of the gene, followed by the expression and purification of the protein.

[0022] Subject of the present invention is a DNA encoding for a thermostable B-type DNA polymerase having a Y-GG/A amino acid motif between the N-terminal 3'-5' exonuclease domain and the C-terminal polymerase domain in the wild type enzyme whereas the tyrosine of this motif is substituted in the mutant enzyme of the polymerase and

whereas this mutant DNA polymerase is suitable for PCR.  
[0023] Preferably, said DNA in wild type form is obtainable from Euryarchaea, more preferably from *Thermococcus aggregans* (*Tag*). Subject of the present invention is also a vector containing the inventive DNA. Suitable vectors are e.g. the following: pET14b/15b/16b/19b (Novagen); pRSET (Invitrogen); pTrcHis (Invitrogen); pHAT10/11/12 (Clontech); pPRO Tet.E/Lar.A (Clontech); pCALn/n-EK (Stratagene); pGEMEX-1/-2 (Promega).

[0024] Furthermore, subjects of the present invention are also cells comprising the above vector. Suitable cells are e.g. *E. coli* BL21, BL21(DE3), BL21(DE3)pLysS, BL21(DE3)pLysE, DH5 $\alpha$ -PRO, JM109 (DE3), TOP10 in combination with the vectors recommended by the suppliers. The gene may have to be subcloned and the protein purification procedure may have to be adapted in the case of different expression vectors.

A sample of the recombinant strain expressing *Tag* DNA polymerase was deposited with the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) Mascheroder Weg 1b, D-38124 Braunschweig (DSM No. 13224).

[0025] A further subject of the invention is the use of the inventive mutant enzymes for synthesizing nucleic acids e.g. in PCR reactions.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

##### Figures:

##### Figure 1.

[0026] Table showing the relative polymerase activities (Pol) and 3'-5'-exonuclease activities (Exo) of *Tag* DNA polymerase and its mutants on double-stranded DNA.

Assays were carried out as described in example 2 and 3, respectively. The activities are expressed as percentage of the activity obtained for the wild-type *Tag* DNA polymerase.

##### Figure 2.

[0027] PCR with *Tag* DNA polymerase mutants.

*Tag* DNA polymerase mutants (1 pmol) were incubated in a 50  $\mu$ l total volume with 10 ng of lambda DNA as template and 30 pmoles of a primer set designed to yield the indicated fragment lengths, 200  $\mu$ M dNTPs and the suitable PCR buffer. Reactions were performed with 10 cycles of 10 sec 94  $^{\circ}$ C, 30 sec 57  $^{\circ}$ C and 3.0 min (A), 4.3 min (B) or 7.0 (C) min of elongation time at 72  $^{\circ}$ C followed by 20 cycles with elongation times increasing by 20 sec/cycle. After the PCR 5  $\mu$ l sample were submitted to electrophoresis on a 1% agarose gel. For the control reaction 2.5 U of *Taq* DNA polymerase, *Pwo* DNA polymerase or Expand<sup>TM</sup> High Fidelity PCR System were used. The labeling of the lanes is described in the legend of figure 6.

##### Figure 3.

[0028] Time-dependent polymerase chain reaction.

1% agarose gels showing 2 kb PCR products from reactions performed with different elongation time (90 sec, 40 sec, 30 sec as indicated) to determine the minimal elongation time.

1 pmol of each *Tag* DNA polymerase mutant or wild type enzyme was added to a mix of 10 ng lambda DNA and primers designed to yield 2 kb DNA fragments. Labeling of the lanes is described in the legend of figure 6. For each mutant duplicate reactions were performed. Right lane of each gel *Pwo*: 2.5 U *Pyrococcus woesei* DNA polymerase (Roche Molecular Biochemicals) as control reaction. Left lane of each gel: Molecular weight marker VI (Roche Molecular Biochemicals). In the 40 sec reaction, mutant GA was omitted. In the 30 sec reaction, for the mutant YS only one reaction was run on the gel.

##### Figure 4.

[0029] 3'-5'-exonuclease processivity.

The *Tag* DNA polymerase mutants tested are indicated on top of the figure. *Tgo* DNA polymerase was used as a control reaction (incubation for 30 sec). Reactions for wild type and mutants of *Tag* DNA polymerase were performed for 4 minutes at 68  $^{\circ}$ C after preincubation for 1 minute at 68 $^{\circ}$ C. Lane "P" is the control reaction (24 mer 5'-DIG-labelled

primer without incubation), lane "-": reaction without heparin (positive control); lane "B": reaction with heparin and  $MnCl_2$  added before addition of the enzyme (negative control); lane "A": heparin and  $MnCl_2$  added after the enzyme (reaction under single turnover conditions).

## 5 Figure 5.

### [0030] Fidelity of *Tag* DNA polymerase mutants.

The fidelity of *Tag* DNA polymerase and its mutants was expressed in relation to the fidelity of *Taq* DNA polymerase. A quotient of 1 means that the polymerase has the same error rate as *Taq* DNA polymerase (mean value  $1.3 \times 10^{-5}$ ). Values >1 reflect the factor by which a polymerase shows less errors than *Taq* DNA polymerase. The bars correspond to mean values calculated from 2-5 independent experiments, error bars missing are smaller than 0.36. Abbreviations for enzymes are as indicated in legend to figure 6. As controls *Pyrococcus woesei* DNA polymerase (Roche Molecular Biochemicals) and Expand High Fidelity PCR System (Roche Molecular Biochemicals) were used ("Taq/Pwo" and "Taq/HiFi").

15

## Figure 6.

### [0031] SDS-PAGE gel analysis of purified mutant proteins.

1  $\mu$ g of each mutant was submitted to electrophoresis on an 10% SDS-PAGE gel. Left: MW, molecular weight marker; WT, *Thermococcus aggregans* wild type DNA polymerase; YF, YW, YS, YN, YH are the corresponding mutants with an exchange at the position of tyrosine 387 to phenylalanine, tryptophan, serine, asparagine, histidine, respectively. GA, mutation of glycine 389 to alanine in the gene of the *Thermococcus aggregans* DNA polymerase. All mutants showed the same chromatographic behaviour and solubility as the wild type enzyme.

## 25 Figure 7.

### [0032] Qualitative exonuclease assay.

A DNA molecular weight marker was used as substrate to test the exonucleolytic activity (DNA molecular weight marker II (MW II), Roche Molecular Biochemicals). 1  $\mu$ g of MW II was incubated for 6 h at 65 °C with 1 pmol of each variant of *Tag* DNA polymerase in the presence (A) or absence (B) of 200  $\mu$ M dNTP. *Tag* mutants are named as explained in legend of figure 6. Exonucleolytic degradation take place only in the absence of deoxynucleotides. The qualitative ranking of the proteins in terms of exonuclease activity is GA > YN > YS > YH > YF = YW = WT.

## Figure 8.

35

### [0033] Consensus sequence motif for B-type DNA polymerases from the order of *Thermococcales* derived from a multiple alignment of amino acid sequences of euryarchaeal and crenarchaeal B-type DNA polymerases.

A region of 24 amino acids containing the Y-GG/A motif was analyzed with the ClustalW Software program (Higgins, EMBL Heidelberg, Germany). In addition to the amino acids conserved in all archaeal B-type DNA polymerases (like the Y-GG/A motif), a consensus sequence "E--RR-R----G(Y)-KE-EE--LWE-" can be defined. This sequence is found in the sequence of all DNA polymerases belonging to the order of the *Thermococcales* and coincides with a homology of >80% of the DNA polymerases.

The sequences of the crenarchaeal species *Sulfolobus solfataricus*, *Sulfolobus acidocaldarius*, *Pyrobaculum islandicum*, *Pyrodicticum occultum*, *Aeropyrum pernix*, *Sulfurisphaera ohwakuensis* and the sequences of several euryarchaeal species *Thermococcus* ("T."), *Pyrococcus* ("P.") and *Methanococcus* ("M.") were aligned.

45

## Figure 9.

### [0034] DNA sequence and deduced amino acid sequence of recombinant wild type *Tag* DNA polymerase.

Three inteins found in the native gene (Acc. No. Y13030) were deleted by PCR (Niehaus, F., Frey, B., Antranikian, A. (1997) *Gene*, **204**, 153-158). Four mutations leading to amino acid exchanges were introduced during PCR. The amino acid exchanges (native→recombinant) are: L3F, A404T, S410C and L482H.

50

## Example 1

55

### Site-directed mutagenesis and expression of *Tag* DNA polymerase mutants

[0035] The cloning of the gene of *Tag* DNA Polymerase (polTY) was described earlier (Niehaus, F., Frey, B., Antran-

ikian, A. (1997) *Gene*, **204**, 153-158). Overexpression of *Tag* DNA Polymerase in *E. coli* was achieved by subcloning its encoding gene into the IPTG-inducible pET15b vector (Novagen) containing an N-terminal His-Tag for purification (the resulting plasmid was named pET15b-TagPol).

The mutants presented in this study were prepared in polymerase chain reactions using primers containing the desired mutations as a mismatch. The forward primer was universally "Kpn-fw", matching to a sequence about 100 bp upstream of the mutation site and contained a *KpnI* restriction site of the *polTy* gene. The reverse primers contained a *SnaBI* restriction site and additionally the desired mutation. The sequences of the oligonucleotides were as follows (mismatch sites for mutagenesis underlined):

SEQ. ID. NO: 1

Kpn-Fw 5'-GCAACCTTGTAGAGTAGAGTGGTACCTGTAAAGGG-3';

SEQ. ID. NO: 2

TagY387F 5'-GCCTCTTTCCGGCTCTTTTACGTATCCTCCCAGGAAAGTAGTCC-3',

SEQ. ID. NO: 3

TagY387H 5'-GCCTCTTTCCGGCTCTTTTACGTATCCTCCCAGGTGAGTAGTCC-3',

SEQ. ID. NO: 4

TagY387N 5'-GCCTCTTTCCGGCTCTTTTACGTATCCTCCCAGGTTAGTAGTCC-3',

SEQ. ID. NO: 5

TagY387S 5'-GCCTCTTTCCGGCTCTTTTACGTATCCTCCCAGGGAAGTAGTCC-3',

SEQ. ID. NO: 6

TagY387W 5'-GCCTCTTTCCGGCTCTTTTACGTATCCTCCCAGCCAAGTAGTCC-3',

SEQ. ID. NO: 7

TagG389A 5'-GCCTCTTTCCGGCTCTTTTACGTATCCAGCCAGGTAAGTAGTCC-3'.

PCR reactions were carried out with Expand™ High Fidelity PCR System (Roche Molecular Biochemicals) using the following program: 2 min 94 °C, 30 cycles of 10 sec 94 °C, 30 sec 55 °C, 30 sec at 72 °C. The resulting 139 bp fragments were digested with the restriction enzymes *KpnI* and *SnaBI* yielding a 101 bp fragment that was ligated into pET15b-TagPol linearized with the restriction enzymes *KpnI* and *SnaBI*. The cloned DNA fragments were sequenced to confirm the presence of the desired mutations.

**[0036]** For protein expression, *E. coli* BL21 (DE3) cells were transformed with the expression vector pET15b-TagPol. Three to five colonies were inoculated in 15 ml of LB medium supplemented with 100 µg Ampicillin per ml and grown to OD<sub>600nm</sub> 0.3. An aliquot (10 ml) of the preculture was used to inoculate 500 ml of LB medium and incubated while shaking at 37°C. At OD<sub>600nm</sub> = 0.6 expression was induced by addition of IPTG (final concentration: 1 mM). After incubation for 3 hours cells were harvested by centrifugation and suspended in 50 mM Tris-HCl/pH 7.5, 10 mM KCl, 0.5 mM EDTA, 4 mM MgCl<sub>2</sub>, 5 mM DTT. Cells were sonicated on ice and the crude extract was heated for 15 min to 80 °C. Cell debris was removed by centrifugation (30 min, 30 000 x g at 4°C).

The supernatant was applied to a Blue Sepharose 3G-A column (Pharmacia) equilibrated with buffer A (50 mM Tris-HCl/ pH 7.5, 10 mM KCl, 4 mM Mg Cl<sub>2</sub>). The protein was eluted with a gradient of 0.01-1.5 M KCl. Active fractions were pooled and dialyzed against 20 mM Tris-HCl/ pH 7.9, 5 mM imidazole, 500 mM NaCl. The sample was applied to a Ni-chelate column (Novagene) equilibrated in the same buffer and eluted with a gradient of 0.005 - 1 M imidazole. Active fractions were pooled and dialyzed against storage buffer (50 mM Tris-HCl/pH 7.5, 100 mM KCl, 0.5 mM EDTA, 5 mM DTT, 50% glycerol). The enzymes were pure as shown by SDS gel electrophoresis (Figure 6).

**Example 2****DNA polymerase assay**

5 [0037] The DNA polymerase activity was determined by measuring the incorporation of  $\alpha$ -( $^{32}$ P)dCTP in a DNA substrate. The test mix (50  $\mu$ l) contained 5  $\mu$ l 10x *Tag* reaction buffer (100 mM Tris-HCl /pH 8.9, 750 mM KCl, 15 MgCl<sub>2</sub>, 100 mM CHAPS), 200  $\mu$ M of each dATP, dGTP, dTTP, 100  $\mu$ M dCTP, 1 mCi  $\alpha$ -( $^{32}$ P)dCTP, 1  $\mu$ g of M13mp9 ssDNA annealed with 0.3  $\mu$ g M13 primer. Assays were performed with 2 and 3  $\mu$ l of enzyme in three different dilutions (final amount of enzyme 2.5 to 15 fmoles) yielding six reactions to calculate a mean value. As a reference *Pwo* DNA polymerase was used. The DNA/primer mix was prepared by heating 277.2  $\mu$ g M13mp9 ssDNA (Roche Molecular Biochemicals) and 156  $\mu$ g M13 sequencing primer (17mer forward primer, Roche Molecular Biochemicals) for 30 minutes to 55 °C and then cooling it for 30 minutes to room temperature.

10 [0038] Assay reactions were incubated for 30 minutes at 65 °C, stopped on ice by addition of 500  $\mu$ l of 10% TCA (4 °C) and kept on ice for another 10 minutes. Samples were filtered over GFC-filter (Whatman), filters washed three times with 5% TCA, dried and submitted to  $\beta$ -counting in 2 ml of scintillation fluid. One unit is defined as the amount of enzyme necessary to incorporate 10 nM dNTP into acid insoluble material at 65 °C in 30 minutes.

**Example 3****Exonuclease Assays****Activity Assay**

25 [0039] 3  $\mu$ l (300 ng) of enzyme (approximately 5 Units of polymerase activity) were incubated with 5 pg of 3H-labelled calf thymus DNA for 4 hours at 65 °C in a buffer containing 10 mM Tris-HCl/pH 8.9, 75 mM KCl, 1.5 mM MgCl<sub>2</sub>, 10 mM CHAPS. Radioactivity liberated from calf thymus DNA was measured in a scintillation counter.

The assay used does not discriminate between the 3'-5' exonuclease activity and the 5'-3' exonuclease activity. 5'-3' exonuclease activity has not been detected phenotypically or genotypically in B-type polymerases of *Thermococcales* (Perler, F. B., Kumar, S. and Kong, H. (1996) *Adv. Prot. Chem.*, **48**, 377-435). Thus the values obtained can be regarded as 3'-5'-exonuclease activity.

30 [0040] In another assay, 1  $\mu$ g of molecular weight marker II (Roche Molecular Biochemicals) were incubated in same buffer as above with 5 U of the indicated protein with or without 200  $\mu$ M dNTP in a final volume of 50  $\mu$ l for 6 hours at 65 °C. The reaction products were separated by electrophoresis on 1% agarose gels.

**3'-5' Exonuclease processivity assay**

35 [0041] The previously described heparin trap method was used (Reddy, M. K., Weitzel, S. E. and von Hippel, P. H. (1992) *J. Biol. Chem.*, **267**(20), 14157-14166; Pisani, F. M., De Felice, M. and Rossi, M. (1998) *Biochemistry*, **37**(42), 15005-15012). A reaction mix (10  $\mu$ l) containing 10 mM Tris-HCl/pH 8.9, 75 mM KCl, 10 mM CHAPS and 0.5 pmoles of a 5'-DIG-labeled 24mer oligonucleotide was prewarmed for 1 minute at 68 °C in a thermocycler. Unless otherwise noted, 1 pmol of the enzyme was preincubated for 1 minute at 68 °C with the substrate. The reaction was started by addition of MnCl<sub>2</sub> (final concentration: 4 mM) and heparin (final concentration: 1 mg/ml) to ensure single turnover conditions. After incubation for 4 minutes, the reaction was stopped by the addition of 5  $\mu$ l of formamide buffer (80% formamide, 10 mM EDTA, 1 mg/ml bromophenol blue, 1 mg/ml xylene xanol). The efficiency of the heparin trap was checked in a control reaction by adding heparin and MnCl<sub>2</sub> prior to the addition of the enzyme. The samples were denatured for 3 minutes at 90 °C and subjected to denaturing gel electrophoresis on a 17.5 % polyacrylamide/8 M urea gel. Gels were blotted to a positively charged nylon membrane (Roche Molecular Biochemicals) and the blots developed with CPD-Star (Roche Molecular Biochemicals) according to the manufacturers instructions.

**Example 4****lacI-based PCR fidelity assay**

50 [0042] We used the *lacI*-based PCR fidelity assay described by Frey and Suppmann (Frey, M. and Suppmann, B. (1995) *Biochemica*, **2**, 34-35). This method is based on the amplification, circularization and transformation of the pUC19 derivative pUCQ17, which contains a functional *lacI*<sup>q</sup> allele (Barnes, W. M. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 2216-2220). PCR-derived mutations in *lacI* result in a de-repression of the expression of *lacZ $\alpha$*  and subsequent formation of a functional  $\beta$ -galactosidase enzyme, which can be easily detected on X-Gal indicator plates.



[0043] The truncated *lacI* gene of pUC19 was substituted by a functional copy of *lacI*<sup>q</sup>. A 178 bp *Pvu II-Afl III* fragment was replaced by a 1121 bp DNA fragment encoding *lacI*<sup>q</sup>. The  $\alpha$ -complementing *E. coli* strain DH5 $\alpha$ , once transformed with the resulting plasmid pUCIQ17 (3632 bp), produces white (LACI<sup>+</sup>) colonies on LB plates containing ampicillin (100  $\mu$ g/ml) and X-Gal (0.004% w/v). For the PCR, pUCIQ17 was linearized by digestion with *Dra II* and used as a template in an amount of 1 or 10 ng. Both primers have *Cla I* cleavage sites at their 5' ends. Oligonucleotide Cla33 (34mer, 24 matches: SEQ. ID. NO: 8: 5'-AGC TTA TCG ATG GCA CTT TTC GGG GAA ATG TGC G-3') and Oligonucleotide Cla55 (36mer, 26 matches: SEQ. ID. NO: 9: 5'- AGC TTA TCG ATA AGC GGA TGC CGG GAG CAG ACA AGC-3') resulted in a PCR product of 3493 bp. The reactions were performed with 1 or 5 pmol of protein in the *Tag* polymerase PCR buffer described below or for the control reactions in the manufacturers PCR buffers with 2.5 U of enzyme. The cycle conditions were 10 sec denaturation at 94 °C, 30 sec annealing at 57 °C and 4 min elongation at 72 °C for 18, 24 or 30 cycles depending on the enzyme.

After PCR, the yield of amplification product was determined at (OD<sub>260nm</sub> or in agarose gel) and the DNA fragments submitted to phenol/chloroform extraction to eliminate any protein. After digestion with *ClaI*, the DNA fragments were purified from a preparative agarose gel. Ligation reactions were carried out with the Rapid Ligation Kit (Roche Molecular Biochemicals), the reactions contained 30 ng DNA. The resulting circular plasmids were transformed in *E. coli* DH5 $\alpha$  as described by Hanahan (Hanahan, D. (1983) *J. Mol. Biol.*, 166, 557-580) and plated on LB Amp/X-Gal plates described above. After incubation overnight at 37 °C, blue and white colonies were counted. The error rate (f) per bp was calculated with a rearranged equation published by Keohavong and Thilly (Keohavong, P. and Thilly, W. G. (1989) *Proc. Natl. Acad.*

*Sci. USA*, 86, 9253-9257):  $f = -\ln F / d \times b$  bp.

[0044] Where F is the fraction of white colonies (white colonies/ total colonies); d is the number of DNA duplications:  $2^d$  = output DNA/input DNA and b is the effective target size (1080 bp) of the *lacI* gene. There are 349 phenotypically identified (by colour screening) single-base substitutions (non-sense and mis-sense) at 179 codons (approximately 50% of the coding region) within the *lacI* gene. (Provost, G. S., Kretz, P. L., Hamner, R. T., Matthews, C. D., Rogers, B. J., Lundberg, K. S., Dyaico, M. J. and Short, J. M. (1993) *Mut. research*, 288, 133-149). Frameshift errors which may occur at every position in the 1080 bp open reading frame of *lacI*, are not taken into account because little information is available for the specific polymerases used in PCR systems except for *Taq* DNA polymerase.

### Example 5

#### Polymerase Chain Reactions

[0045] PCR was performed in a buffer optimized for *Tag* DNA polymerase and its mutants: 10 mM Tris-HCl/pH 8.9, 75 mM KCl, 1.5 MgCl<sub>2</sub>, 10 mM CHAPS, 200  $\mu$ M dNTP. 10 ng of  $\lambda$  DNA were used as a template and 30 pmol of each primer (20 bp, designed to yield the products of the desired length):

Lambda 1, universal forward primer: 5'-GAT GAG TTC GTG TCC GTA CAA CA-3', SEQ. ID. NO: 10,

Lambda 3.3: 5'-CTC ATC AGC AGA TCA TCT TCA GG-3', SEQ. ID. NO: 11,

Lambda 8: 5'-ACT CCA GCG TCT CAT CTT TAT GC-3', SEQ. ID. NO: 12,

Lambda 9: 5'-GAT GGT GAT CCT CTC TCG TTT GC-3', SEQ. ID. NO: 13.

Lambda 3.3, 8 and 9 were used as reverse primers for the amplification of 3.3 kb, 5 kb and 7.5 kb fragments respectively. Template, primers and nucleotides were prepared in mix 1 in a volume of 25  $\mu$ l. Then 25  $\mu$ l of mix 2 containing the buffer and enzyme (1 pmol *Tag* wild type or mutant; or 2.5 U control enzyme) were added. All reactions were prepared in duplicate. The amplification was performed in a 2400 GeneAmp thermocycler (Perkin Elmer). The cycle conditions were: 2 min at 94 °C, 10 cycles with 10 sec denaturation at 94 °C, 30 sec annealing at 58 °C and elongation at 72 °C. [0046] Elongation times depended on the length of the product (3 min for 3.3 kb, 4.3 min for 5 kb and 7 min for 7.5 kb). Another 20 cycles were performed with increasing the elongation times by 20 sec/cycle. The reaction was finished by 7 minutes at 72 °C. The tubes were kept at 4°C until separation by electrophoresis on 1% agarose gels.

## EP 1 132 474 A1

[0047] The improvement of the PCR performance of the mutants was also studied in a "time-dependent PCR". A 2 kb fragment was amplified from lambda DNA as described above. In these studies the elongation time of the PCR was stepwise reduced (90 sec, 40 sec, 30 sec) to determine for each enzyme the minimal elongation time that was sufficient to amplify the 2 kb fragment. The following primers were used:

Lambda 1, universal forward primer: 5'-GAT GAG TTC GTG TCC GTA CAA CA-3', SEQ. ID. NO: 14,  
Lambda 6, reverse primer: 5'-CTT CAT CAT CGA GAT AGC TGT CG-3', SEQ. ID. NO: 15.

The temperature profile was as described above. The elongation times were kept constant over 30 cycles.

SEQUENCE LISTING

5 <110> Roche Diagnostics GmbH

<120> Mutant B-type DNA polymerase exhibiting improved performance in PCR

10 <130> 5328/00/EP

<140> 00105115.6

<141> 2000-03-11

15 <160> 34

<170> PatentIn Ver. 2.1

<210> 1

20 <211> 35

<212> DNA

<213> Artificial Sequence

<220>

25 <223> Description of Artificial Sequence: Artificial

<400> 1

gcaaccttgt agagtagagt ggtacctgtt aaggg 35

30 <210> 2

<211> 44

<212> DNA

<213> Artificial Sequence

35 <220>

<223> Description of Artificial Sequence: Artificial

<400> 2

gcctctttcc ggctctttta cgtatcctcc caggaaagta gtcc 44

40 <210> 3

<211> 44

<212> DNA

<213> Artificial Sequence

45 <220>

<223> Description of Artificial Sequence: Artificial

<400> 3

50 gcctctttcc ggctctttta cgtatcctcc caggtagagta gtcc 44

<210> 4

<211> 44

<212> DNA

55 <213> Artificial Sequence

EP 1 132 474 A1

<220>  
 <223> Description of Artificial Sequence: Artificial  
 5 <400> 4  
 gcctctttcc ggctctttta cgtatcctcc caggtagta gtcc 44  
 <210> 5  
 10 <211> 44  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 15 <223> Description of Artificial Sequence: Artificial  
 <400> 5  
 gcctctttcc ggctctttta cgtatcctcc cagggaagta gtcc 44  
 <210> 6  
 20 <211> 44  
 <212> DNA  
 <213> Artificial Sequence  
 25 <220>  
 <223> Description of Artificial Sequence: Artificial  
 <400> 6  
 30 gcctctttcc ggctctttta cgtatcctcc cagccaagta gtcc 44  
 <210> 7  
 <211> 44  
 <212> DNA  
 <213> Artificial Sequence  
 35 <220>  
 <223> Description of Artificial Sequence: Artificial  
 <400> 7  
 40 gcctctttcc ggctctttta cgtatccagc caggtaagta gtcc 44  
 <210> 8  
 <211> 34  
 <212> DNA  
 45 <213> Artificial Sequence  
 <220>  
 <223> Description of Artificial Sequence: Artificial  
 50 <400> 8  
 agcttatcga tggcactttt cggggaaatg tgcg 34  
 55

<210> 9  
 <211> 36  
 <212> DNA  
 5 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: Artificial  
  
 10 <400> 9  
 agcttatcga taagcggatg ccgggagcag acaagc 36  
  
 <210> 10  
 <211> 23  
 15 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: Artificial  
 20  
 <400> 10  
 gatgagttcg tgtccgtaca aca 23  
  
 <210> 11  
 25 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 30 <223> Description of Artificial Sequence: Artificial  
  
 <400> 11  
 ctcatcagca gatcatcttc agg 23  
  
 35 <210> 12  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 40 <220>  
 <223> Description of Artificial Sequence: Artificial  
  
 <400> 12  
 45 actccagcgt ctcatcttta tgc 23  
  
 <210> 13  
 <211> 23  
 <212> DNA  
 50 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: Artificial  
  
 55

# EP 1 132 474 A1

<400> 13  
 gatggtgatc ctctctcggt tgc 23  
 5 <210> 14  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 10 <220>  
 <223> Description of Artificial Sequence: Artificial  
 <400> 14  
 15 gatgagttcg tgtccgtaca aca 23  
 <210> 15  
 <211> 23  
 <212> DNA  
 20 <213> Artificial Sequence  
 <220>  
 <223> Description of Artificial Sequence: Artificial  
 25 <400> 15  
 cttcatcatc gagatagctg tcg 23  
 <210> 16  
 <211> 25  
 30 <212> PRT  
 <213> T. aggregans  
 <400> 16  
 Glu Tyr Arg Arg Arg Leu Arg Thr Thr Tyr Leu Gly Gly Tyr Val Lys  
 35 1 5 10 15  
 Glu Pro Glu Arg Gly Leu Trp Glu Asn  
 20 25  
 40 <210> 17  
 <211> 25  
 <212> PRT  
 <213> T. litoralis  
 45 <400> 17  
 Glu Tyr Lys Arg Arg Leu Arg Thr Thr Tyr Leu Gly Gly Tyr Val Lys  
 1 5 10 15  
 50 Glu Pro Glu Lys Gly Leu Trp Glu Asn  
 20 25  
 55

EP 1 132 474 A1

<210> 18  
 <211> 25  
 <212> PRT  
 5 <213> T. fumicolans  
  
 <400> 18  
 Glu Leu Glu Arg Arg Xaa Arg Gly Gly Tyr Ala Gly Gly Tyr Val Lys  
     1                    5                    10                    15  
 10 Glu Pro Glu Arg Gly Leu Trp Glu Asn  
                     20                    25  
  
 15 <210> 19  
 <211> 25  
 <212> PRT  
 <213> T. spec. 9N7  
  
 20 <400> 19  
 Glu Leu Ala Arg Arg Xaa Arg Gly Gly Tyr Ala Gly Gly Tyr Val Lys  
     1                    5                    10                    15  
 25 Glu Arg Glu Arg Gly Leu Trp Glu Asn  
                     20                    25  
  
 30 <210> 20  
 <211> 25  
 <212> PRT  
 <213> T. gorgonarius  
  
 <400> 20  
 35 Glu Leu Ala Arg Arg Xaa Arg Glu Ser Tyr Ala Gly Gly Tyr Val Lys  
     1                    5                    10                    15  
 Glu Pro Glu Arg Gly Leu Trp Glu Asn  
                     20                    25  
  
 40 <210> 21  
 <211> 25  
 <212> PRT  
 45 <213> P. spec. KOD  
  
 <400> 21  
 Glu Leu Ala Arg Arg Xaa Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys  
     1                    5                    10                    15  
 50 Glu Pro Glu Arg Gly Leu Trp Glu Asn  
                     20                    25  
  
 55

EP 1 132 474 A1

<210> 22  
 <211> 25  
 <212> PRT  
 5 <213> P. abyssii  
  
 <400> 22  
 Glu Tyr Glu Arg Arg Leu Arg Glu Ser Tyr Glu Gly Gly Tyr Val Lys  
     1                    5                    10                    15  
 10 Glu Pro Glu Lys Gly Leu Trp Glu Asn  
                     20                    25  
  
 15 <210> 23  
 <211> 25  
 <212> PRT  
 <213> P. furiosus  
  
 20 <400> 23  
 Glu Tyr Gln Arg Arg Leu Arg Glu Ser Tyr Thr Gly Gly Phe Val Lys  
     1                    5                    10                    15  
 25 Glu Pro Glu Lys Gly Leu Trp Glu Asn  
                     20                    25  
  
 30 <210> 24  
 <211> 25  
 <212> PRT  
 <213> P. horikoshii  
  
 35 <400> 24  
 Glu Tyr Glu Arg Arg Leu Arg Glu Ser Tyr Glu Gly Gly Tyr Val Lys  
     1                    5                    10                    15  
 Glu Pro Glu Lys Gly Leu Trp Glu Asn  
                     20                    25  
  
 40 <210> 25  
 <211> 25  
 <212> PRT  
 <213> M. jannaschii  
  
 45 <400> 25  
 Glu Tyr Arg Arg Arg Val Leu Thr Thr Tyr Glu Gly Gly Tyr Val Lys  
     1                    5                    10                    15  
 50 Glu Pro Glu Lys Gly Met Phe Glu Asp  
                     20                    25  
  
 55



EP 1 132 474 A1

<210> 26  
 <211> 25  
 <212> PRT  
 5 <213> M. voltae  
  
 <400> 26  
 Ser Tyr Arg Glu Arg Ala Lys Phe Ser Tyr Glu Gly Gly Tyr Val Arg  
 1 5 10 15  
 10 Glu Pro Leu Lys Gly Ile Gln Glu Asn  
 20 25  
  
 15 <210> 27  
 <211> 25  
 <212> PRT  
 <213> S. solifataricus  
  
 20 <400> 27  
 Thr Ser Ala Leu Ile Lys Gly Lys Gly Tyr Lys Gly Ala Val Val Ile  
 1 5 10 15  
 25 Asp Pro Pro Ala Gly Ile Phe Phe Asn  
 20 25  
  
 30 <210> 28  
 <211> 25  
 <212> PRT  
 <213> S. acidocaldarius  
  
 35 <400> 28  
 Thr Ala Ala Val Ile Lys Gly Lys Lys Tyr Lys Gly Ala Val Val Ile  
 1 5 10 15  
 Asp Pro Pro Ala Gly Val Tyr Phe Asn  
 20 25  
  
 40 <210> 29  
 <211> 25  
 <212> PRT  
 <213> P. islandicum  
  
 45 <400> 29  
 Thr Lys Ala Ile Ile Lys Gly Lys Lys Tyr Ala Gly Ala Val Val Leu  
 1 5 10 15  
 50 Asp Pro Pro Leu Gly Ile Phe Phe Asn  
 20 25  
  
 55

EP 1 132 474 A1

<210> 30  
 <211> 25  
 <212> PRT  
 5 <213> P. occultum  
  
 <400> 30  
 Ser Glu Ala Leu Ile Lys Gly Lys Lys Tyr Gln Gly Ala Leu Val Leu  
     1                    5                    10                    15  
 10 Asp Pro Pro Ser Gly Ile Tyr Phe Asn  
                     20                    25  
  
 15 <210> 31  
 <211> 25  
 <212> PRT  
 <213> A. pernix  
  
 20 <400> 31  
 Val Gly Ala Ile Ile Lys Asp Lys Lys Tyr Arg Gly Ala Ile Val Leu  
     1                    5                    10                    15  
 25 Asp Pro Pro Val Gly Ile Phe Phe Arg  
                     20                    25  
  
 30 <210> 32  
 <211> 25  
 <212> PRT  
 <213> S. chwakuensis  
  
 35 <400> 32  
 Thr Ala Ala Ile Ser Lys Gly Lys Arg Tyr Lys Gly Ala Val Val Ile  
     1                    5                    10                    15  
 Asp Pro Pro Ala Gly Val Phe Phe Asn  
                     20                    25  
  
 40 <210> 33  
 <211> 2325  
 <212> DNA  
 <213> T. aggregans  
  
 45 <220>  
 <221> CDS  
 <222> (1)..(2325)  
  
 50 <400> 33  
 atg ata ttt gac act gac tac ata aca aag gac ggt aaa ccc ata att 48  
 Met Ile Phe Asp Thr Asp Tyr Ile Thr Lys Asp Gly Lys Pro Ile Ile  
     1                    5                    10                    15  
  
 55

EP 1 132 474 A1

	cga att ttc aag aaa gag aac ggg gaa ttt aaa ata gaa ctt gat cca	96
	Arg Ile Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Leu Asp Pro	
	20 25 30	
5	cat ttt cag ccc tac att tac gct ctt ctc aaa gat gac tcc gct att	144
	His Phe Gln Pro Tyr Ile Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile	
	35 40 45	
10	gat gaa ata aaa gca ata aaa ggc gag aga cac gga aaa att gtg aga	192
	Asp Glu Ile Lys Ala Ile Lys Gly Glu Arg His Gly Lys Ile Val Arg	
	50 55 60	
15	gta gtc gat gca gtg aaa gtc aag aag aaa ttt ttg ggg aga gat gtt	240
	Val Val Asp Ala Val Lys Val Lys Lys Lys Phe Leu Gly Arg Asp Val	
	65 70 75 80	
20	gag gtc tgg aag ctt ata ttt gag cat ccc caa gac gtc ccg gcc cta	288
	Glu Val Trp Lys Leu Ile Phe Glu His Pro Gln Asp Val Pro Ala Leu	
	85 90 95	
25	agg ggc aag ata agg gaa cat cca gct gtg att gac att tat gag tac	336
	Arg Gly Lys Ile Arg Glu His Pro Ala Val Ile Asp Ile Tyr Glu Tyr	
	100 105 110	
30	gac ata ccc ttt gcc aag cgc tac ctc ata gac aag ggc ttg atc cct	384
	Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile Pro	
	115 120 125	
35	atg gag ggc gac gag gag ctt aag cta atg gcc ttc gac att gag acg	432
	Met Glu Gly Asp Glu Glu Leu Lys Leu Met Ala Phe Asp Ile Glu Thr	
	130 135 140	
40	ttt tac cac gag gga gac gag ttt ggg aag ggc gag ata ata atg ata	480
	Phe Tyr His Glu Gly Asp Glu Phe Gly Lys Gly Glu Ile Ile Met Ile	
	145 150 155 160	
45	agc tac gcc gat gag gaa gag gca agg gta att aca tgg aag aat att	528
	Ser Tyr Ala Asp Glu Glu Glu Ala Arg Val Ile Thr Trp Lys Asn Ile	
	165 170 175	
50	gat ctg ccc tac gtt gat gtt gta tcc aac gaa agg gag atg ata aag	576
	Asp Leu Pro Tyr Val Asp Val Val Ser Asn Glu Arg Glu Met Ile Lys	
	180 185 190	
55	cgg ttt gtg caa att gtc agg gaa aaa gac ccg gat gtc ctg ata act	624
	Arg Phe Val Gln Ile Val Arg Glu Lys Asp Pro Asp Val Leu Ile Thr	
	195 200 205	
60	tac aat gga gac aac ttt gat ttg ccg tac ctt ata aaa agg gca gag	672
	Tyr Asn Gly Asp Asn Phe Asp Leu Pro Tyr Leu Ile Lys Arg Ala Glu	
	210 215 220	

EP 1 132 474 A1

	aag tta gga gtt act ctt ctc ttg ggg agg gac aaa gaa cac ccc gag	720
	Lys Leu Gly Val Thr Leu Leu Leu Gly Arg Asp Lys Glu His Pro Glu	
	225 230 235 240	
5	ccc aag att cac aga atg ggc gat agc ttt gcc gtg gaa att aaa ggc	768
	Pro Lys Ile His Arg Met Gly Asp Ser Phe Ala Val Glu Ile Lys Gly	
	245 250 255	
10	aga att cac ttt gat ctc ttc ccg gtt gtg cgg aga acc ata aac ctc	816
	Arg Ile His Phe Asp Leu Phe Pro Val Val Arg Arg Thr Ile Asn Leu	
	260 265 270	
15	cca aca tac acg ctt gag gca gtt tat gaa gcc gtc ttg gga aaa acc	864
	Pro Thr Tyr Thr Leu Glu Ala Val Tyr Glu Ala Val Leu Gly Lys Thr	
	275 280 285	
20	aaa agc aag ctg ggt gcg gag gaa atc gcc gcc atc tgg gaa aca gag	912
	Lys Ser Lys Leu Gly Ala Glu Glu Ile Ala Ala Ile Trp Glu Thr Glu	
	290 295 300	
25	gag agc atg aag aag ctg gcc cag tac tcg atg gaa gat gct agg gca	960
	Glu Ser Met Lys Lys Leu Ala Gln Tyr Ser Met Glu Asp Ala Arg Ala	
	305 310 315 320	
30	act tat gaa ctc gga aaa gag ttt ttc ccc atg gag gca gag cta gca	1008
	Thr Tyr Glu Leu Gly Lys Glu Phe Phe Pro Met Glu Ala Glu Leu Ala	
	325 330 335	
35	aag cta ata ggc caa agc gta tgg gac gtc tca aga tca agc act ggc	1056
	Lys Leu Ile Gly Gln Ser Val Trp Asp Val Ser Arg Ser Ser Thr Gly	
	340 345 350	
40	aac ctt gta gag tgg tac ctg tta agg gtg gca tat gag agg aat gag	1104
	Asn Leu Val Glu Trp Tyr Leu Leu Arg Val Ala Tyr Glu Arg Asn Glu	
	355 360 365	
45	ctc gct ccg aac aag ccg gat gaa gaa gag tac aga agg cgt tta agg	1152
	Leu Ala Pro Asn Lys Pro Asp Glu Glu Glu Tyr Arg Arg Arg Leu Arg	
	370 375 380	
50	act act tac ctg gga gga tac gta aaa gag ccg gaa aga ggc tta tgg	1200
	Thr Thr Tyr Leu Gly Gly Tyr Val Lys Glu Pro Glu Arg Gly Leu Trp	
	385 390 395 400	
55	gag aac atc acc tat tta gac ttt agg tgc cta tac ccc tca att ata	1248
	Glu Asn Ile Thr Tyr Leu Asp Phe Arg Cys Leu Tyr Pro Ser Ile Ile	
	405 410 415	
60	gtt acc cac aac gtc tcc cct gac act tta gaa aga gaa ggc tgc aag	1296
	Val Thr His Asn Val Ser Pro Asp Thr Leu Glu Arg Glu Gly Cys Lys	
	420 425 430	

EP 1 132 474 A1

	aat tac gat gtt gcc ccg ata gta ggt tat aag ttc tgc aag gat ttt	1344
	Asn Tyr Asp Val Ala Pro Ile Val Gly Tyr Lys Phe Cys Lys Asp Phe	
	435 440 445	
5	ccc ggt ttc att cca tct ata ctc ggg gaa tta atc aca atg agg caa	1392
	Pro Gly Phe Ile Pro Ser Ile Leu Gly Glu Leu Ile Thr Met Arg Gln	
	450 455 460	
10	gaa ata aag aag aag atg aaa gct aca att gac cca ata gaa aag aaa	1440
	Glu Ile Lys Lys Lys Met Lys Ala Thr Ile Asp Pro Ile Glu Lys Lys	
	465 470 475 480	
15	atg ctt gat tat agg caa aga gct gtt aaa ttg cac gca aac agc tat	1488
	Met Leu Asp Tyr Arg Gln Arg Ala Val Lys Leu His Ala Asn Ser Tyr	
	485 490 495	
20	tac ggt tat atg ggc tat ccc aag gcg agg tgg tac tcg aag gaa tgt	1536
	Tyr Gly Tyr Met Gly Tyr Pro Lys Ala Arg Trp Tyr Ser Lys Glu Cys	
	500 505 510	
25	gcc gaa agc gtt acc gcg tgg gga agg cac tac ata gaa atg acc ata	1584
	Ala Glu Ser Val Thr Ala Trp Gly Arg His Tyr Ile Glu Met Thr Ile	
	515 520 525	
30	aaa gag ata gag gag aaa ttt gga ttt aag gtg cta tat gcc gac act	1632
	Lys Glu Ile Glu Glu Lys Phe Gly Phe Lys Val Leu Tyr Ala Asp Thr	
	530 535 540	
35	gat ggt ttt tac gcc aca ata ccg gga gaa aaa cct gaa aca atc aaa	1680
	Asp Gly Phe Tyr Ala Thr Ile Pro Gly Glu Lys Pro Glu Thr Ile Lys	
	545 550 555 560	
40	aag aaa gct aag gaa ttc tta aaa tac ata aac tcc aaa ctt ccc ggt	1728
	Lys Lys Ala Lys Glu Phe Leu Lys Tyr Ile Asn Ser Lys Leu Pro Gly	
	565 570 575	
45	ctg ctc gag ctt gag tat gag ggc ttt tac ttg aga gga ttt ttc gtc	1776
	Leu Leu Glu Leu Glu Tyr Glu Gly Phe Tyr Leu Arg Gly Phe Phe Val	
	580 585 590	
50	gca aag aag cgc tat gcg gtt ata gac gaa gaa ggt agg ata acg aca	1824
	Ala Lys Lys Arg Tyr Ala Val Ile Asp Glu Glu Gly Arg Ile Thr Thr	
	595 600 605	
55	agg ggt ctg gaa gtt gta agg agg gac tgg agc gaa ata gcc aaa gag	1872
	Arg Gly Leu Glu Val Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu	
	610 615 620	
60	acc cag gct aaa gtc ttg gag gca ata ctt aaa gaa gat agt gtc gaa	1920
	Thr Gln Ala Lys Val Leu Glu Ala Ile Leu Lys Glu Asp Ser Val Glu	
	625 630 635 640	

EP 1 132 474 A1

5      aaa gct gtg gaa atc gtt aag gac gtt gtt gag gag ata gca aaa tac      1968  
       Lys Ala Val Glu Ile Val Lys Asp Val Val Glu Glu Ile Ala Lys Tyr  
                                 645                                  650                                  655  
  
       caa gtc ccg ctt gaa aag ctt gtt atc cac gag cag att acc aag gat      2016  
       Gln Val Pro Leu Glu Lys Leu Val Ile His Glu Gln Ile Thr Lys Asp  
                                 660                                  665                                  670  
  
 10     cta agt gaa tac aaa gcc att ggg cct cat gta gca ata gca aag agg      2064  
       Leu Ser Glu Tyr Lys Ala Ile Gly Pro His Val Ala Ile Ala Lys Arg  
                                 675                                  680                                  685  
  
 15     ctt gct gca aag gga ata aaa gtg aga ccc ggc acg ata ata agc tat      2112  
       Leu Ala Ala Lys Gly Ile Lys Val Arg Pro Gly Thr Ile Ile Ser Tyr  
                                 690                                  695                                  700  
  
 20     atc gtc ctc agg gga agc gga aag ata agt gac agg gta att ttg ctt      2160  
       Ile Val Leu Arg Gly Ser Gly Lys Ile Ser Asp Arg Val Ile Leu Leu  
       705                                  710                                  715                                  720  
  
       tca gag tat gat ccg aaa aaa cac aag tac gac ccc gac tac tac ata      2208  
       Ser Glu Tyr Asp Pro Lys Lys His Lys Tyr Asp Pro Asp Tyr Tyr Ile  
                                 725                                  730                                  735  
  
 25     gaa aac caa gtt ctg ccg gcg gtg ctt agg atc ctt gaa gcc ttc ggc      2256  
       Glu Asn Gln Val Leu Pro Ala Val Leu Arg Ile Leu Glu Ala Phe Gly  
                                 740                                  745                                  750  
  
 30     tac aga aaa gag gac tta aaa tac caa agc tca aaa cag gtt gga ctg      2304  
       Tyr Arg Lys Glu Asp Leu Lys Tyr Gln Ser Ser Lys Gln Val Gly Leu  
                                 755                                  760                                  765  
  
 35     gac gcg tgg ctt aag aag tag      2325  
       Asp Ala Trp Leu Lys Lys  
                                 770                                  775  
  
 40     <210> 34  
       <211> 775  
       <212> PRT  
       <213> T. aggregans  
  
       <400> 34  
 45     Met Ile Phe Asp Thr Asp Tyr Ile Thr Lys Asp Gly Lys Pro Ile Ile  
       1                                  5                                  10                                  15  
  
       Arg Ile Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Leu Asp Pro  
                                 20                                  25                                  30  
  
 50     His Phe Gln Pro Tyr Ile Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile  
                                 35                                  40                                  45  
  
 55

EP 1 132 474 A1

	Asp	Glu	Ile	Lys	Ala	Ile	Lys	Gly	Glu	Arg	His	Gly	Lys	Ile	Val	Arg	
	50						55					60					
5	Val	Val	Asp	Ala	Val	Lys	Val	Lys	Lys	Lys	Phe	Leu	Gly	Arg	Asp	Val	
	65					70					75					80	
	Glu	Val	Trp	Lys	Leu	Ile	Phe	Glu	His	Pro	Gln	Asp	Val	Pro	Ala	Leu	
					85					90					95		
10	Arg	Gly	Lys	Ile	Arg	Glu	His	Pro	Ala	Val	Ile	Asp	Ile	Tyr	Glu	Tyr	
				100					105					110			
	Asp	Ile	Pro	Phe	Ala	Lys	Arg	Tyr	Leu	Ile	Asp	Lys	Gly	Leu	Ile	Pro	
15			115					120					125				
	Met	Glu	Gly	Asp	Glu	Glu	Leu	Lys	Leu	Met	Ala	Phe	Asp	Ile	Glu	Thr	
	130						135					140					
20	Phe	Tyr	His	Glu	Gly	Asp	Glu	Phe	Gly	Lys	Gly	Glu	Ile	Ile	Met	Ile	
	145					150					155					160	
	Ser	Tyr	Ala	Asp	Glu	Glu	Glu	Ala	Arg	Val	Ile	Thr	Trp	Lys	Asn	Ile	
					165					170					175		
25	Asp	Leu	Pro	Tyr	Val	Asp	Val	Val	Ser	Asn	Glu	Arg	Glu	Met	Ile	Lys	
				180						185				190			
	Arg	Phe	Val	Gln	Ile	Val	Arg	Glu	Lys	Asp	Pro	Asp	Val	Leu	Ile	Thr	
30			195					200					205				
	Tyr	Asn	Gly	Asp	Asn	Phe	Asp	Leu	Pro	Tyr	Leu	Ile	Lys	Arg	Ala	Glu	
	210						215					220					
35	Lys	Leu	Gly	Val	Thr	Leu	Leu	Leu	Gly	Arg	Asp	Lys	Glu	His	Pro	Glu	
	225					230					235					240	
	Pro	Lys	Ile	His	Arg	Met	Gly	Asp	Ser	Phe	Ala	Val	Glu	Ile	Lys	Gly	
					245					250					255		
40	Arg	Ile	His	Phe	Asp	Leu	Phe	Pro	Val	Val	Arg	Arg	Thr	Ile	Asn	Leu	
				260					265					270			
	Pro	Thr	Tyr	Thr	Leu	Glu	Ala	Val	Tyr	Glu	Ala	Val	Leu	Gly	Lys	Thr	
45			275					280					285				
	Lys	Ser	Lys	Leu	Gly	Ala	Glu	Glu	Ile	Ala	Ala	Ile	Trp	Glu	Thr	Glu	
	290						295					300					
50	Glu	Ser	Met	Lys	Lys	Leu	Ala	Gln	Tyr	Ser	Met	Glu	Asp	Ala	Arg	Ala	
	305					310					315					320	
	Thr	Tyr	Glu	Leu	Gly	Lys	Glu	Phe	Phe	Pro	Met	Glu	Ala	Glu	Leu	Ala	
					325					330					335		
55																	

EP 1 132 474 A1

Lys Leu Ile Gly Gln Ser Val Trp Asp Val Ser Arg Ser Ser Thr Gly  
 340 345 350  
 5 Asn Leu Val Glu Trp Tyr Leu Leu Arg Val Ala Tyr Glu Arg Asn Glu  
 355 360 365  
 Leu Ala Pro Asn Lys Pro Asp Glu Glu Glu Tyr Arg Arg Arg Leu Arg  
 370 375 380  
 10 Thr Thr Tyr Leu Gly Gly Tyr Val Lys Glu Pro Glu Arg Gly Leu Trp  
 385 390 395 400  
 Glu Asn Ile Thr Tyr Leu Asp Phe Arg Cys Leu Tyr Pro Ser Ile Ile  
 15 405 410 415  
 Val Thr His Asn Val Ser Pro Asp Thr Leu Glu Arg Glu Gly Cys Lys  
 420 425 430  
 20 Asn Tyr Asp Val Ala Pro Ile Val Gly Tyr Lys Phe Cys Lys Asp Phe  
 435 440 445  
 Pro Gly Phe Ile Pro Ser Ile Leu Gly Glu Leu Ile Thr Met Arg Gln  
 450 455 460  
 25 Glu Ile Lys Lys Lys Met Lys Ala Thr Ile Asp Pro Ile Glu Lys Lys  
 465 470 475 480  
 Met Leu Asp Tyr Arg Gln Arg Ala Val Lys Leu His Ala Asn Ser Tyr  
 30 485 490 495  
 Tyr Gly Tyr Met Gly Tyr Pro Lys Ala Arg Trp Tyr Ser Lys Glu Cys  
 500 505 510  
 35 Ala Glu Ser Val Thr Ala Trp Gly Arg His Tyr Ile Glu Met Thr Ile  
 515 520 525  
 Lys Glu Ile Glu Glu Lys Phe Gly Phe Lys Val Leu Tyr Ala Asp Thr  
 530 535 540  
 40 Asp Gly Phe Tyr Ala Thr Ile Pro Gly Glu Lys Pro Glu Thr Ile Lys  
 545 550 555 560  
 Lys Lys Ala Lys Glu Phe Leu Lys Tyr Ile Asn Ser Lys Leu Pro Gly  
 565 570 575  
 45 Leu Leu Glu Leu Glu Tyr Glu Gly Phe Tyr Leu Arg Gly Phe Phe Val  
 580 585 590  
 Ala Lys Lys Arg Tyr Ala Val Ile Asp Glu Glu Gly Arg Ile Thr Thr  
 595 600 605  
 50 Arg Gly Leu Glu Val Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu  
 610 615 620

55



# EP 1 132 474 A1

Thr Gln Ala Lys Val Leu Glu Ala Ile Leu Lys Glu Asp Ser Val Glu  
 625 630 635 640  
 5 Lys Ala Val Glu Ile Val Lys Asp Val Val Glu Glu Ile Ala Lys Tyr  
 645 650 655  
 Gln Val Pro Leu Glu Lys Leu Val Ile His Glu Gln Ile Thr Lys Asp  
 660 665 670  
 10 Leu Ser Glu Tyr Lys Ala Ile Gly Pro His Val Ala Ile Ala Lys Arg  
 675 680 685  
 Leu Ala Ala Lys Gly Ile Lys Val Arg Pro Gly Thr Ile Ile Ser Tyr  
 690 695 700  
 15 Ile Val Leu Arg Gly Ser Gly Lys Ile Ser Asp Arg Val Ile Leu Leu  
 705 710 715 720  
 20 Ser Glu Tyr Asp Pro Lys Lys His Lys Tyr Asp Pro Asp Tyr Tyr Ile  
 725 730 735  
 Glu Asn Gln Val Leu Pro Ala Val Leu Arg Ile Leu Glu Ala Phe Gly  
 740 745 750  
 25 Tyr Arg Lys Glu Asp Leu Lys Tyr Gln Ser Ser Lys Gln Val Gly Leu  
 755 760 765  
 30 Asp Ala Trp Leu Lys Lys  
 770 775

## 35 Claims

1. A thermostable mutant B-type DNA polymerase having a Y-GG/A amino acid motif between the N-terminal 3'-5'-exonuclease domain and the C-terminal polymerase domain in the wild-type form of the polymerase whereas preferably the tyrosine of this motif is substituted in the mutant form of the polymerase and whereas this mutant DNA polymerase is suitable for polymerase chain reactions.
2. A mutant B-type DNA polymerase according to claim 1 having an amino acid with an aromatic side chain at the position of the tyrosine.
3. A mutant thermostable B-type DNA polymerase according to claim 1 having a Y→F, Y→W or Y→H mutation.
4. A mutant B-type DNA polymerase according to claim 1 having an amino acid with an hydrophilic side chain at the position of the tyrosine.
5. A mutant thermostable B-type DNA polymerase according to claim 1 having a Y→N or Y→S mutation.
6. A mutant thermostable B-type DNA polymerase according to claims 1-5 whereas the wild type form is obtainable from Euryarchaea.
7. A mutant thermostable B-type DNA polymerase according to claims 1-6 whereas the wild type form is obtainable from *Thermococcus aggregans*.
8. A mutant of a thermostable B-type DNA polymerase according to claims 1-6 whereas the amino acid sequence

## EP 1 132 474 A1

of the wild type form is  $\geq 80\%$  homologue to the amino acid sequence of wild type *Taq* DNA polymerase.

9. A DNA encoding a thermostable mutant DNA polymerase according to claims 1-8.

5 10. A vector containing a DNA according to claim 9.

11. A transformed host cell comprising a vector according to claim 10.

10 12. A process for obtaining a polymerase according to claim 1-8 comprising the steps of cloning and mutagenesis of the gene, followed by the expression and purification of the protein.

13. Use of a polymerase according to claim 1-8 for synthesizing nucleic acids.

15 14. Use of a polymerase according to claims 1-8 for PCR reactions.

20

25

30

35

40

45

50

55

Figure 1

Pol activity and Exo activity in %

	WT	Y387F	Y387W	Y387H	Y387N	Y387S	G389A
Pol	100	160	92	93,6	6,4	17,8	10,7
Exo	100	90	71	98	205	187	236
Pol/Exo	1	1,77	1,29	0,96	0,03	0,09	0,04

Figure 2

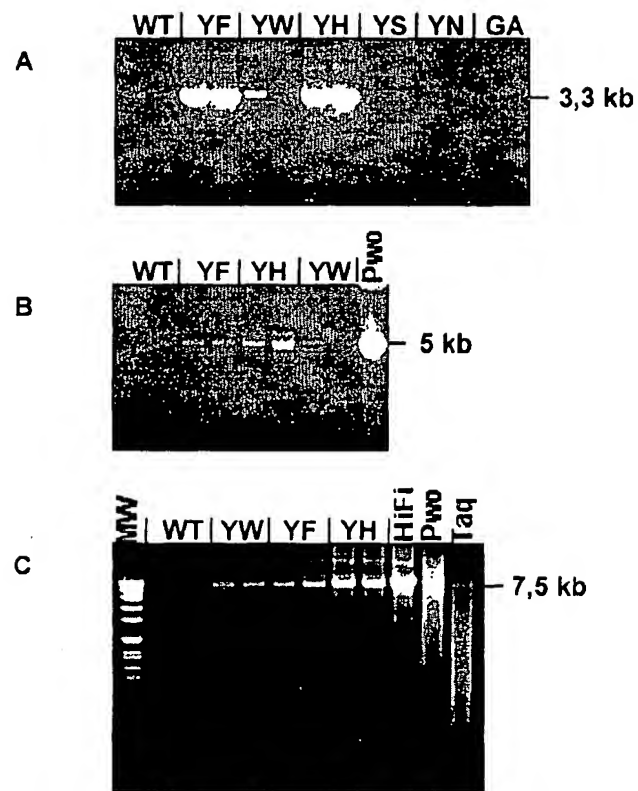


Figure 3

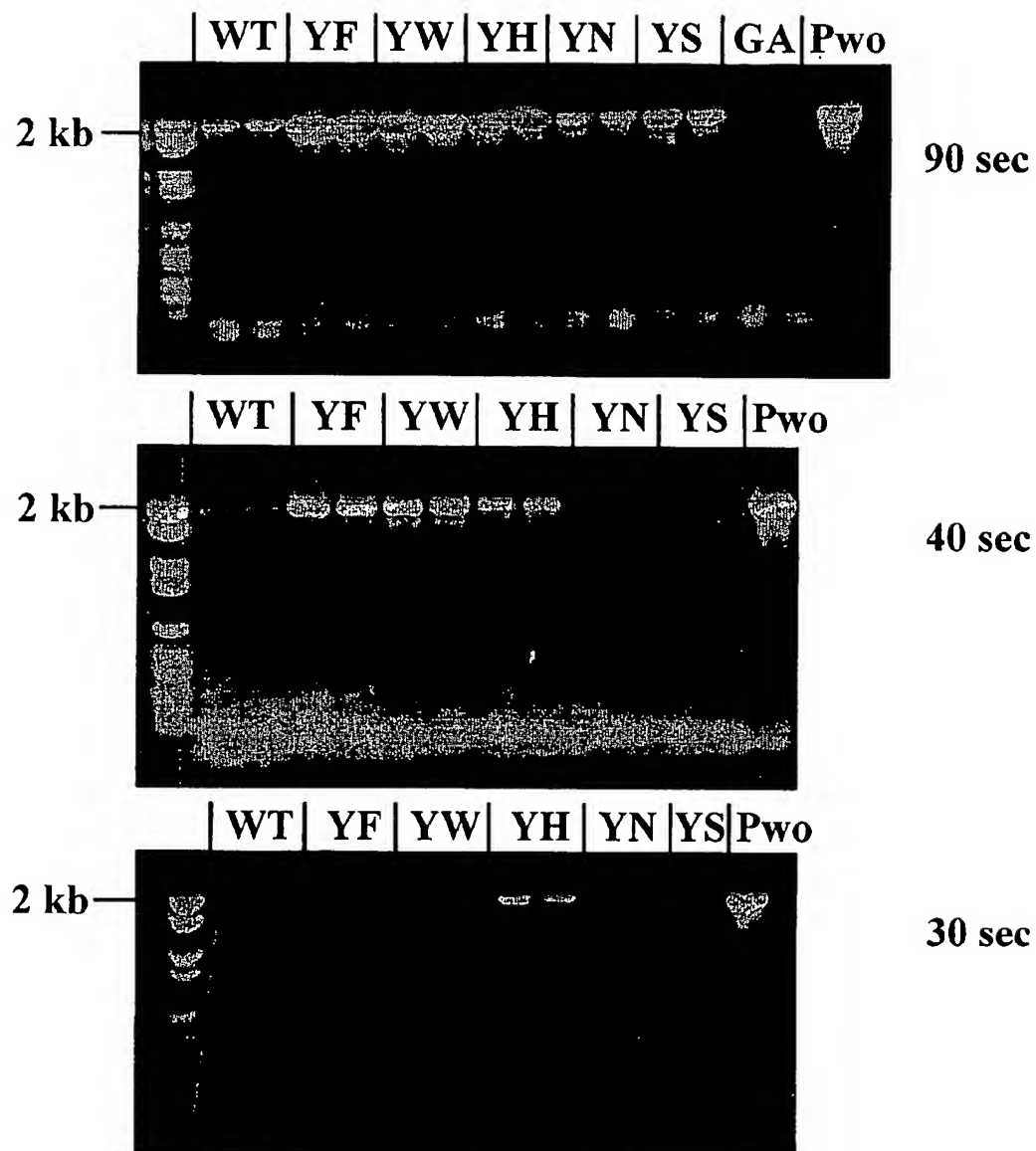


Figure 4

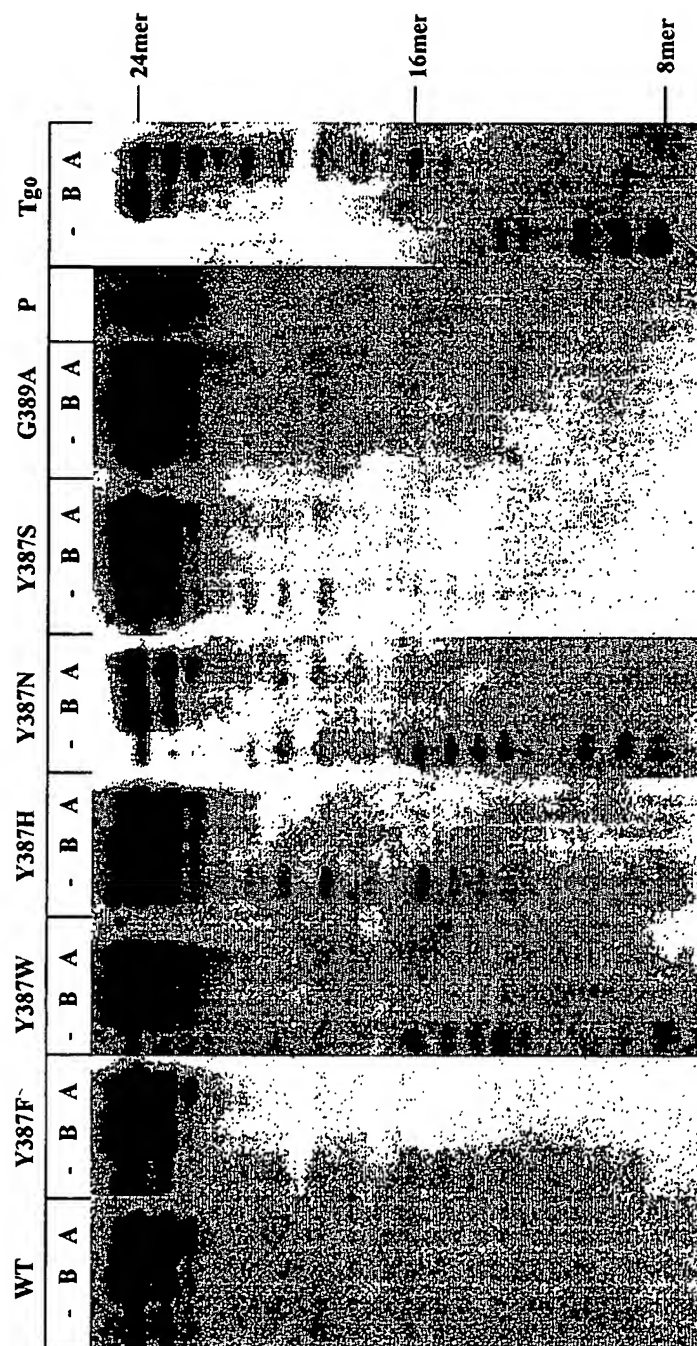


Figure 5

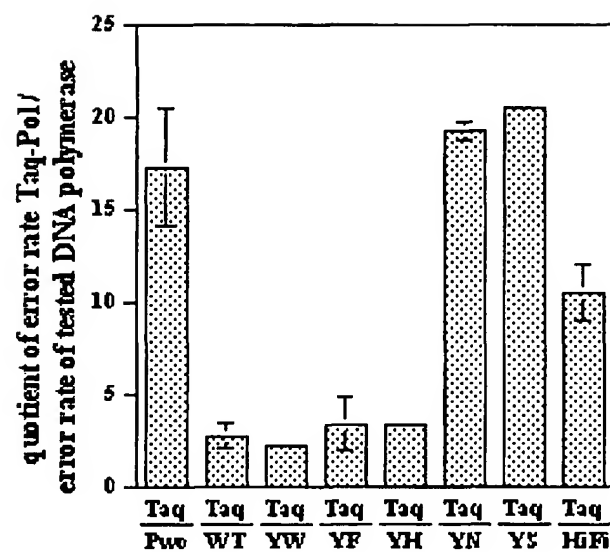


Figure 6

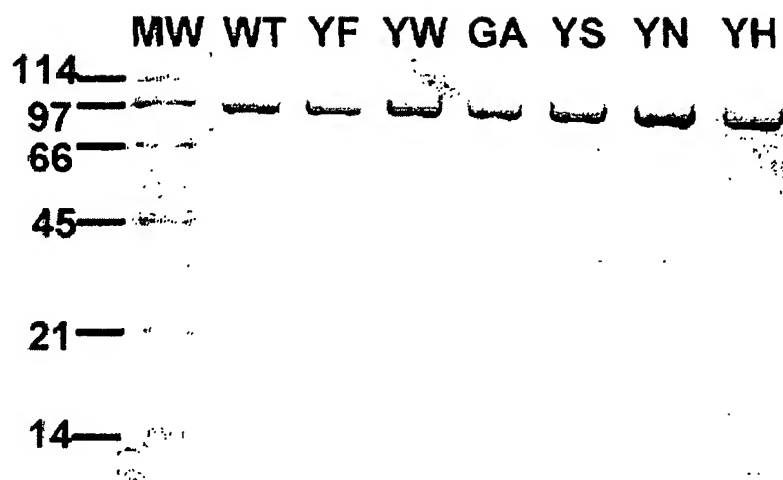
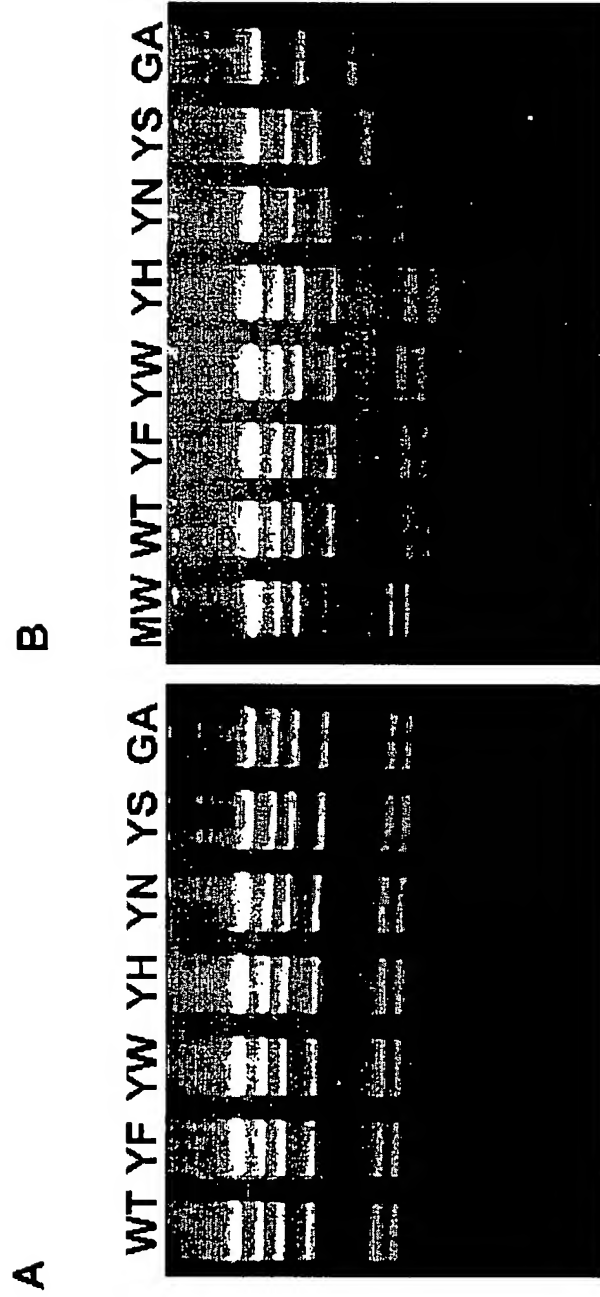




Figure 7



	E	R	R	R		G (Y)	K	E	E	L	W	E	consensus													
<i>T. aggregans</i>	E	Y	R	R	L	R	T	T	Y	L	G	G	Y	V	K	E	P	E	R	G	L	W	E	N		
<i>T. litoralis</i>	E	Y	K	R	R	L	R	T	T	Y	L	G	G	Y	V	K	E	P	E	R	G	L	W	E	N	
<i>T. tumidclans</i>	E	L	E	R	R	-	R	G	G	Y	A	G	G	Y	V	K	E	P	E	R	G	L	W	E	N	
<i>T. spec. 9N7</i>	E	L	A	R	R	-	R	G	G	Y	A	G	G	Y	V	K	E	P	E	R	G	L	W	E	N	
<i>T. gorgonarius</i>	E	L	A	R	R	-	R	E	S	Y	A	G	G	Y	V	K	E	P	E	R	G	L	W	E	N	
<i>P. spec. KOD</i>	E	L	A	R	R	-	R	O	S	Y	E	G	G	Y	V	K	E	P	E	R	G	L	W	E	N	
<i>P. abyssii</i>	E	Y	E	R	R	L	R	E	S	Y	E	G	G	Y	V	K	E	P	E	R	G	L	W	E	N	
<i>P. furiosus</i>	E	Y	Q	R	R	L	R	E	S	Y	T	G	G	F	V	K	E	P	E	R	G	L	W	E	N	
<i>P. horikoshii</i>	E	Y	E	R	R	L	R	E	S	Y	E	G	G	Y	V	K	E	P	E	R	G	L	W	E	N	
<i>M. jannaschii</i>	E	Y	R	R	B	V	L	T	T	Y	E	G	G	Y	V	K	E	P	E	R	G	L	W	E	N	
<i>M. voltae</i>	S	Y	R	E	R	A	K	E	S	Y	E	G	G	Y	V	R	E	P	L	K	G	I	O	E	N	
<i>S. solfataricus</i>	T	S	A	L	I	K	G	K	G	Y	K	G	A	V	V	I	D	P	P	A	G	I	F	F	N	
<i>S. acidocaldarius</i>	T	A	A	V	I	K	G	K	K	Y	K	G	A	V	V	I	D	P	P	A	G	V	Y	F	N	
<i>P. islandicum</i>	T	K	A	I	I	K	G	K	K	Y	A	G	A	V	V	L	D	P	P	L	G	I	F	F	N	
<i>P. occultum</i>	S	E	A	L	I	K	G	K	K	Y	O	G	A	L	V	L	D	P	P	S	G	I	Y	F	N	
<i>A. pernix</i>	V	G	A	I	I	K	D	K	K	Y	R	G	A	I	V	L	D	P	P	V	G	I	F	F	N	
<i>S. chwalkuensis</i>	T	A	A	I	S	K	G	K	R	Y	K	G	A	V	V	I	D	P	P	A	G	V	F	F	N	

SEQ. ID. NO: 16  
 SEQ. ID. NO: 17  
 SEQ. ID. NO: 18  
 SEQ. ID. NO: 19  
 SEQ. ID. NO: 20  
 SEQ. ID. NO: 21  
 SEQ. ID. NO: 22  
 SEQ. ID. NO: 23  
 SEQ. ID. NO: 24  
 SEQ. ID. NO: 25  
 SEQ. ID. NO: 26  
 SEQ. ID. NO: 27  
 SEQ. ID. NO: 28  
 SEQ. ID. NO: 29  
 SEQ. ID. NO: 30  
 SEQ. ID. NO: 31  
 SEQ. ID. NO: 32

Figure 8

Figure 9/1

SEQ. ID. NO: 33

SEQ. ID. NO: 34

1/1

ATG ATA TTT GAC ACT GAC TAC ATA ACA AAG GAC GGT AAA CCC ATA ATT CGA ATT TTC AAG  
 Met ile phe asp thr asp tyr ile thr lys asp gly lys pro ile ile arg ile phe lys  
 61/21  
 AAA GAG AAC GGG GAA TTT AAA ATA GAA CTT GAT CCA CAT TTT CAG CCC TAC ATT TAC GCT  
 lys glu asn gly glu phe lys ile glu leu asp pro his phe gln pro tyr ile tyr ala  
 121/41  
 CTT CTC AAA GAT GAC TCC GCT ATT GAT GAA ATA AAA GCA ATA AAA GGC GAG AGA CAC GGA  
 leu leu lys asp asp ser ala ile asp glu ile lys ala ile lys gly glu arg his gly  
 181/61  
 AAA ATT GTG AGA GTA GTC GAT GCA GTG AAA GTC AAG AAG AAA TTT TTG GGG AGA GAT GTT  
 lys ile val arg val val asp ala val lys val lys lys lys phe leu gly arg asp val  
 241/81  
 GAG GTC TGG AAG CTT ATA TTT GAG CAT CCC CAA GAC GTC CCG GCC CTA AGG GGC AAG ATA  
 glu val trp lys leu ile phe glu his pro gln asp val pro ala leu arg gly lys ile  
 301/101  
 AGG GAA CAT CCA GCT GTG ATT GAC ATT TAT GAG TAC GAC ATA CCC TTT GCC AAG CGC TAC  
 arg glu his pro ala val ile asp ile tyr glu tyr asp ile pro phe ala lys arg tyr  
 361/121  
 CTC ATA GAC AAG GGC TTG ATC CCT ATG GAG GGC GAC GAG GAG CTT AAG CTA ATG GCC TTC  
 leu ile asp lys gly leu ile pro met glu gly asp glu glu leu lys leu met ala phe  
 421/141  
 GAC ATT GAG ACG TTT TAC CAC GAG GGA GAC GAG TTT GGG AAG GGC GAG ATA ATA ATG ATA  
 asp ile glu thr phe tyr his glu gly asp glu phe gly lys gly glu ile ile met ile  
 481/161  
 AGC TAC GCC GAT GAG GAA GAG GCA AGG GTA ATT ACA TGG AAG AAT ATT GAT CTG CCC TAC  
 ser tyr ala asp glu glu glu ala arg val ile thr trp lys asn ile asp leu pro tyr  
 541/181  
 GTT GAT GTT GTA TCC AAC GAA AGG GAG ATG ATA AAG CGG TTT GTG CAA ATT GTC AGG GAA  
 val asp val val ser asn glu arg glu met ile lys arg phe val gln ile val arg glu  
 601/201  
 AAA GAC CCG GAT GTC CTG ATA ACT TAC AAT GGA GAC AAC TTT GAT TTG CCG TAC CTT ATA  
 lys asp pro asp val leu ile thr tyr asn gly asp asn phe asp leu pro tyr leu ile  
 661/221  
 AAA AGG GCA GAG AAG TTA GGA GTT ACT CTT CTC TTG GGG AGG GAC AAA GAA CAC CCC GAG  
 lys arg ala glu lys leu gly val thr leu leu leu gly arg asp lys glu his pro glu  
 721/241  
 CCC AAG ATT CAC AGA ATG GGC GAT AGC TTT GCC GTG GAA ATT AAA GGC AGA ATT CAC TTT  
 pro lys ile his arg met gly asp ser phe ala val glu ile lys gly arg ile his phe  
 781/261  
 GAT CTC TTC CCG GTT GTG CGG AGA ACC ATA AAC CTC CCA ACA TAC ACG CTT GAG GCA GTT  
 asp leu phe pro val val arg arg thr ile asn leu pro thr tyr thr leu glu ala val  
 841/281  
 TAT GAA GCC GTC TTG GGA AAA ACC AAA AGC AAG CTG GGT GCG GAG GAA ATC GCC GCC ATC  
 tyr glu ala val leu gly lys thr lys ser lys leu gly ala glu glu ile ala ala ile  
 901/301  
 TGG GAA ACA GAG GAG AGC ATG AAG AAG CTG GCC CAG TAC TCG ATG GAA GAT GCT AGG GCA  
 trp glu thr glu glu ser met lys lys leu ala gln tyr ser met glu asp ala arg ala  
 961/321  
 ACT TAT GAA CTC GGA AAA GAG TTT TTC CCC ATG GAG GCA GAG CTA GCA AAG CTA ATA GGC  
 thr tyr glu leu gly lys glu phe phe pro met glu ala glu leu ala lys leu ile gly  
 1021/341  
 CAA AGC GTA TGG GAC GTC TCA AGA TCA AGC ACT GGC AAC CTT GTA GAG TGG TAC CTG TTA  
 gln ser val trp asp val ser arg ser ser thr gly asn leu val glu trp tyr leu leu  
 1081/361

Figure 9/2

SEQ. ID. NO: 33  
 SEQ. ID. NO: 34

AGG GTG GCA TAT GAG AGG AAT GAG CTC GCT CCG AAC AAG CCG GAT GAA GAA GAG TAC AGA  
 arg val ala tyr glu arg asn glu leu ala pro asn lys pro asp glu glu glu tyr arg  
 1141/381  
 AGG CGT TTA AGG ACT ACT TAC CTG GGA GGA TAC GTA AAA GAG CCG GAA AGA GGC TTA TGG  
 arg arg leu arg thr thr tyr leu gly gly tyr val lys glu pro glu arg gly leu trp  
 1201/401  
 GAG AAC ATC ACC TAT TTA GAC TTT AGG TGC CTA TAC CCC TCA ATT ATA GTT ACC CAC AAC  
 glu asn ile thr tyr leu asp phe arg cys leu tyr pro ser ile ile val thr his asn  
 1261/421  
 GTC TCC CCT GAC ACT TTA GAA AGA GAA GGC TGC AAG AAT TAC GAT GTT GCC CCG ATA GTA  
 val ser pro asp thr leu glu arg glu gly cys lys asn tyr asp val ala pro ile val  
 1321/441  
 GGT TAT AAG TTC TGC AAG GAT TTT CCC GGT TTC ATT CCA TCT ATA CTC GGG GAA TTA ATC  
 gly tyr lys phe cys lys asp phe pro gly phe ile pro ser ile leu gly glu leu ile  
 1381/461  
 ACA ATG AGG CAA GAA ATA AAG AAG AAG ATG AAA GCT ACA ATT GAC CCA ATA GAA AAG AAA  
 thr met arg gln glu ile lys lys lys met lys ala thr ile asp pro ile glu lys lys  
 1441/481  
 ATG CTT GAT TAT AGG CAA AGA GCT GTT AAA TTG CAC GCA AAC AGC TAT TAC GGT TAT ATG  
 met leu asp tyr arg gln arg ala val lys leu his ala asn ser tyr tyr gly tyr met  
 1501/501  
 GGC TAT CCC AAG GCG AGG TGG TAC TCG AAG GAA TGT GCC GAA AGC GTT ACC GCG TGG GGA  
 gly tyr pro lys ala arg trp tyr ser lys glu cys ala glu ser val thr ala trp gly  
 1561/521  
 AGG CAC TAC ATA GAA ATG ACC ATA AAA GAG ATA GAG GAG AAA TTT GGA TTT AAG GTG CTA  
 arg his tyr ile glu met thr ile lys glu ile glu glu lys phe gly phe lys val leu  
 1621/541  
 TAT GCC GAC ACT GAT GGT TTT TAC GCC ACA ATA CCG GGA GAA AAA CCT GAA ACA ATC AAA  
 tyr ala asp thr asp gly phe tyr ala thr ile pro gly glu lys pro glu thr ile lys  
 1681/561  
 AAG AAA GCT AAG GAA TTC TTA AAA TAC ATA AAC TCC AAA CTT CCC GGT CTG CTC GAG CTT  
 lys lys ala lys glu phe leu lys tyr ile asn ser lys leu pro gly leu leu glu leu  
 1741/581  
 GAG TAT GAG GGC TTT TAC TTG AGA GGA TTT TTC GTC GCA AAG AAG CGC TAT GCG GTT ATA  
 glu tyr glu gly phe tyr leu arg gly phe phe val ala lys lys arg tyr ala val ile  
 1801/601  
 GAC GAA GAA GGT AGG ATA ACG ACA AGG GGT CTG GAA GTT GTA AGG AGG GAC TGG AGC GAA  
 asp glu glu gly arg ile thr thr arg gly leu glu val val arg arg asp trp ser glu  
 1861/621  
 ATA GCC AAA GAG ACC CAG GCT AAA GTC TTG GAG GCA ATA CTT AAA GAA GAT AGT GTC GAA  
 ile ala lys glu thr gln ala lys val leu glu ala ile leu lys glu asp ser val glu  
 1921/641  
 AAA GCT GTG GAA ATC GTT AAG GAC GTT GTT GAG GAG ATA GCA AAA TAC CAA GTC CCG CTT  
 lys ala val glu ile val lys asp val val glu glu ile ala lys tyr gin val pro leu  
 1981/661  
 GAA AAG CTT GTT ATC CAC GAG CAG ATT ACC AAG GAT CTA AGT GAA TAC AAA GCC ATT GGG  
 glu lys leu val ile his glu gln ile thr lys asp leu ser glu tyr lys ala ile gly  
 2041/681  
 CCT CAT GTA GCA ATA GCA AAG AGG CTT GCT GCA AAG GGA ATA AAA GTG AGA CCC GGC ACG  
 pro his val ala ile ala lys arg leu ala ala lys gly ile lys val arg pro gly thr  
 2101/701  
 ATA ATA AGC TAT ATC GTC CTC AGG GGA AGC GGA AAG ATA AGT GAC AGG GTA ATT TTG CTT  
 ile ile ser tyr ile val leu arg gly ser gly lys ile ser asp arg val ile leu leu  
 2161/721

Figure 9/3

SEQ. ID. NO: 33

SEQ. ID. NO: 34

TCA GAG TAT GAT CCG AAA AAA CAC AAG TAC GAC CCC GAC TAC TAC ATA GAA AAC CAA GTT  
ser glu tyr asp pro lys lys his lys tyr asp pro asp tyr tyr ile glu asn gln val  
2221/741

CTG CCG GCG GTG CTT AGG ATC CTT GAA GCC TTC GGC TAC AGA AAA GAG GAC TTA AAA TAC  
leu pro ala val leu arg ile leu glu ala phe gly tyr arg lys glu asp leu lys tyr  
2281/761

CAA AGC TCA AAA CAG GTT GGA CTG GAC GCG TGG CTT AAG AAG TAG  
gln ser ser lys gln val gly leu asp ala trp leu lys lys AMB



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 01 10 4583

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
D,X	PISANI, F.M. ET AL.: "Amino Acid Residues Involved in Determining the Processivity of the 3'-5' Exonuclease Activity in a Family B DNA Polymerase from the Thermoacidophilic Archaeon Sulfolobus solfataricus" BIOCHEMISTRY, vol. 37, no. 42, 20 October 1998 (1998-10-20), pages 15005-15012, XP002145197	1-5,9-13	C12N15/52 C12N9/00
A	* abstract * * page 15006, column 1, line 21 - line 33 * * page 15009, column 2, line 12 - line 28 * * page 15010; figure 4A * * page 15010, column 1, line 16 - line 24 * * page 15011; tables 1,2 * * page 15011, column 2, line 3 - line 36 * * page 15012, column 1, line 6 - line 18 *	14	
D,A	TRUNIGER, V. ET AL.: "A DNA binding motif coordinating synthesis and degradation in proofreading DNA polymerases" EMBO JOURNAL, vol. 15, no. 13, 1 July 1996 (1996-07-01), pages 3430-3441, XP000925854 * the whole document *	1-14	
A	DE 196 11 759 A (BOEHRINGER MANNHEIM GMBH) 2 October 1997 (1997-10-02) * the whole document *	1-14	
-/--			
The present search report has been drawn up for all claims			
Place of search BERLIN		Date of completion of the search 11 July 2001	Examiner Fuchs, U
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date O : document cited in the application L : document cited for other reasons &amp; : member of the same patent family, corresponding document</p>			

EPO FORM 1503 (01/02) (P04001)



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 01 10 4583

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
P,X	<p>BÖHLKE, K. ET AL.: "PCR performance of the B-type DNA polymerase from the thermophilic euryarchaeon Thermococcus aggregans improved by mutations in the Y-GG/A motif"</p> <p>NUCLEIC ACIDS RESEARCH, vol. 28, no. 20, 15 October 2000 (2000-10-15), pages 3910-3917, XP002171733</p> <p>* the whole document *</p> <p>-----</p>	1-14	
			TECHNICAL FIELDS SEARCHED (Int.Cl.7)
The present search report has been drawn up for all claims			
Place of search BERLIN		Date of completion of the search 11 July 2001	Examiner Fuchs, U
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone</p> <p>Y : particularly relevant if combined with another document of the same category</p> <p>A : technological background</p> <p>O : non-written disclosure</p> <p>P : intermediate document</p>		<p>T : theory or principle underlying the invention</p> <p>E : earlier patent document, but published on, or after the filing date</p> <p>D : document cited in the application</p> <p>L : document cited for other reasons</p> <p>&amp; : member of the same patent family, corresponding document</p>	

EPO FORM 1503 03 82 (P/C/C01)

EP 01 10 4583

11-07-2001

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
DE 19611759 A	02-10-1997	WO 9735988 A	02-10-1997
-----			

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82